aws

User Guide

Amazon Macie

Amazon Macie: User Guide

Table of Contents

What is Amazon Macie?

Amazon Macie is a data security service that discovers sensitive data by using machine learning and pattern matching, provides visibility into data security risks, and enables automated protection against those risks.

To help you manage the security posture of your organization's Amazon Simple Storage Service (Amazon S3) data estate, Macie provides you with an inventory of your S3 general purpose buckets, and automatically evaluates and monitors the buckets for security and access control. If Macie detects a potential issue with the security or privacy of your data, such as a bucket that becomes publicly accessible, Macie generates a finding for you to review and remediate as necessary.

Macie also automates discovery and reporting of sensitive data to provide you with a better understanding of the data that your organization stores in Amazon S3. To detect sensitive data, you can use built-in criteria and techniques that Macie provides, custom criteria that you define, or a combination of the two. If Macie detects sensitive data in an S3 object, Macie generates a finding to notify you of the sensitive data that it found.

In addition to findings, Macie provides statistics and information that offer insight into the security posture of your Amazon S3 data and where sensitive data might reside in your data estate. The statistics and information can guide your decisions to perform deeper investigations of specific S3 buckets and objects. You can review and analyze findings, statistics, and other information by using the Amazon Macie console or the Amazon Macie API. You can also leverage Macie integration with Amazon EventBridge and AWS Security Hub to monitor, process, and remediate findings by using other services, applications, and systems.

Topics

Features of Macie

Here are some of the key ways that Amazon Macie can help you discover, monitor, and protect your sensitive data in Amazon S3.

Automate the discovery of sensitive data

With Macie, you can automate discovery and reporting of sensitive data in two ways: by configuring Macie to perform automated sensitive data discovery, and by creating and running sensitive data discovery jobs. If Macie detects sensitive data in an S3 object, it creates a sensitive data finding for you. The finding provides a detailed report of the sensitive data that Macie detected.

Automated sensitive data discovery provides broad visibility into where sensitive data might reside in your Amazon S3 data estate. With this option, Macie continually evaluates your S3 bucket inventory and uses sampling techniques to identify and select representative S3 objects from your buckets. Macie then retrieves and analyzes the selected objects, inspecting them for sensitive data.

Sensitive data discovery jobs provide deeper, more targeted analysis. With this option, you define the breadth and depth of the analysis—the S3 buckets to analyze, the sampling depth, and custom criteria that derive from properties of S3 objects. You can also configure a job to run only once for on-demand analysis and assessment, or on a recurring basis for periodic analysis, assessment, and monitoring.

Both options can help you build and maintain a comprehensive view of the data that your organization stores in Amazon S3 and any security or compliance risks for that data.

Discover a variety of sensitive data types

To discover sensitive data with Macie, you can use built-in criteria and techniques, such as machine learning and pattern matching, to analyze objects in S3 buckets. These criteria and techniques, referred to as managed data identifiers, can detect a large and growing list of sensitive data types for many countries and regions, including multiple types of personally identifiable information (PII), financial information, and credentials data.

You can also use custom data identifiers. A custom data identifier is a set of criteria that you define to detect sensitive data—a regular expression (*regex*) that defines a text pattern to match and, optionally, character sequences and a proximity rule that refine the results. With this type of identifier, you can detect sensitive data that reflects your particular scenarios, intellectual property, or proprietary data. You can supplement the managed data identifiers that Macie provides.

To fine tune analyses, you can also use allow lists. Allow lists define specific text and text patterns that you want Macie to ignore in S3 objects. These are typically sensitive data

exceptions for your particular scenarios or environment—for example, the names of public representatives for your organization, public phone numbers for your organization, or sample data that your organization uses for testing.

Evaluate and monitor data for security and access control

When you enable Macie, Macie automatically generates and begins maintaining an inventory of your S3 general purpose buckets. Macie also begins evaluating and monitoring the buckets for security and access control. If Macie detects a potential issue with the security or privacy of a bucket, it creates a policy finding for you.

In addition to findings, a dashboard gives you a snapshot of aggregated statistics for your Amazon S3 data. This includes statistics for key metrics such as the number of buckets that are publicly accessible or shared with other AWS accounts. You can drill down on each statistic to review the supporting data.

Macie also provides detailed information and statistics for individual S3 buckets in your inventory. The data includes breakdowns of a bucket's public access and encryption settings, and the size and number of objects that Macie can analyze to detect sensitive data in the bucket. You can browse the inventory, or sort and filter the inventory by certain fields.

Review and analyze findings

In Macie, a finding is a detailed report of sensitive data that Macie detected in an S3 object or a potential issue with the security or privacy of an S3 general purpose bucket. Each finding provides a severity rating, information about the affected resource, and additional details, such as when and how Macie detected the data or issue.

To review, analyze, and manage findings, you can use **Findings** pages on the Amazon Macie console. These pages list your findings and provide the details of individual findings. They also provide multiple options for grouping, filtering, sorting, and suppressing findings. You can also use the Amazon Macie API to retrieve and review findings. If you use the API, you can pass the data to another application, service, or system for deeper analysis, long-term storage, or reporting.

Monitor and process findings with other services and systems

To support integration with other services and systems, Macie publishes findings to Amazon EventBridge as events. EventBridge is a serverless event bus service that can route findings data to targets such as AWS Lambda functions and Amazon Simple Notification Service (Amazon SNS) topics. With EventBridge, you can monitor and process findings in near real time as part of your existing security and compliance workflows.

You can configure Macie to also publish findings to AWS Security Hub. Security Hub is a service that provides a comprehensive view of your security posture across your AWS environment and helps you check your environment against security industry standards and best practices. With Security Hub, you can more easily evaluate and process findings as part of a broader analysis of your organization's security posture in AWS. You can also aggregate findings from multiple AWS Regions, and then evaluate and process aggregated findings data from a single Region.

Centrally manage multiple Macie accounts

If your AWS environment has multiple accounts, you can centrally manage Macie for accounts in your environment. You can do this in two ways, by integrating Macie with AWS Organizations or by sending and accepting membership invitations in Macie.

In a multiple-account configuration, a designated Macie administrator can perform certain tasks and access certain Macie settings, data, and resources for accounts that are members of the same organization. Tasks include reviewing information about S3 buckets that are owned by member accounts, reviewing policy findings for those buckets, and inspecting the buckets for sensitive data. If the accounts are associated through AWS Organizations, the Macie administrator can also enable Macie for member accounts in the organization.

Develop and manage resources programmatically

In addition to the Amazon Macie console, you can interact with Macie by using the Amazon Macie API. The Amazon Macie API gives you comprehensive, programmatic access to your Macie settings, data, and resources.

To interact with Macie programmatically, you can send HTTPS requests directly to Macie or use a current version of an AWS command line tool or an AWS SDK. AWS provides tools and SDKs that consist of libraries and sample code for various languages and platforms, such as PowerShell, Java, Go, Python, C++, and .NET.

Accessing Macie

Amazon Macie is available in most AWS Regions. For a list of Regions where Macie is currently available, see Amazon Macie endpoints and quotas in the *AWS General Reference*. For information about managing AWS Regions for your AWS account, see Enable or disable AWS Regions in your account in the *AWS Account Management Reference Guide*.

In each Region, you can work with Macie in any of the following ways.

AWS Management Console

The AWS Management Console is a browser-based interface that you can use to create and manage AWS resources. As part of that console, the Amazon Macie console provides access to your Macie account, data, and resources. You can perform any Macie task by using the Macie console—review statistics and other information about your S3 buckets, create and run sensitive data discovery jobs, review and analyze findings, and more.

AWS command line tools

With AWS command line tools, you can issue commands at your system's command line to perform Macie tasks and AWS tasks. Using the command line can be faster and more convenient than using the console. The command line tools are also useful if you want to build scripts that perform tasks.

AWS provides two sets of command line tools: the AWS Command Line Interface (AWS CLI) and the AWS Tools for PowerShell. For information about installing and using the AWS CLI, see the AWS Command Line Interface User Guide. For information about installing and using the Tools for PowerShell, see the AWS Tools for PowerShell User Guide.

AWS SDKs

AWS provides SDKs that consist of libraries and sample code for various programming languages and platforms—for example, Java, Go, Python, C++, and .NET. The SDKs provide convenient, programmatic access to Macie and other AWS services. They also handle tasks such as cryptographically signing requests, managing errors, and retrying requests automatically. For information about installing and using the AWS SDKs, see Tools to Build on AWS.

Amazon Macie REST API

The Amazon Macie REST API gives you comprehensive, programmatic access to your Macie account, data, and resources. With this API, you can send HTTPS requests directly to Macie. However, unlike the AWS command line tools and SDKs, use of this API requires your application to handle low-level details such as generating a hash to sign a request. For information about this API, see the Amazon Macie API Reference.

Pricing for Macie

As with other AWS products, there are no contracts or minimum commitments for using Amazon Macie.

Macie pricing is based on several dimensions—evaluating and monitoring S3 buckets for security and access control, monitoring S3 objects for automated sensitive data discovery, and analyzing S3 objects to discover and report sensitive data in the objects. For more information, see Amazon Macie pricing.

To help you understand and forecast the cost of using Macie, Macie provides estimated usage costs for your account. You can review these estimates on the Amazon Macie console and access them with the Amazon Macie API. Depending on how you use the service, you might incur additional costs for using other AWS services in combination with certain Macie features, such as retrieving bucket data from Amazon S3 and using customer managed AWS KMS keys to decrypt objects for analysis.

When you enable Macie for the first time, your AWS account is automatically enrolled in the 30-day free trial of Macie. This includes individual accounts that are enabled as part of an organization in AWS Organizations. During the free trial, there's no charge for using Macie in the applicable AWS Region to evaluate and monitor your S3 buckets for security and access control. Depending on your account settings, the free trial can also include performing automated sensitive data discovery for your Amazon S3 data. The free trial doesn't include running sensitive data discovery jobs to discover and report sensitive data in S3 objects.

To help you understand and forecast the cost of using Macie after the free trial ends, Macie provides you with estimated usage costs based on your use of Macie during the trial. Your usage data also indicates the amount of time that remains before your free trial ends. You can review this data on the Amazon Macie console and access it with the Amazon Macie API. For more information, see Participating in the free trial.

Related services

To further secure your data, workloads, and applications in AWS, consider using the following AWS services in combination with Amazon Macie.

AWS Security Hub

AWS Security Hub gives you a comprehensive view of the security state of your AWS resources and helps you check your AWS environment against security industry standards and best practices. It does this partly by consuming, aggregating, organizing, and prioritizing your security findings from multiple AWS services (including Macie) and supported AWS Partner Network (APN) products. Security Hub helps you analyze your security trends and identify the highest priority security issues across your AWS environment.

To learn more about Security Hub, see the AWS Security Hub User Guide. To learn about using Macie and Security Hub together, see Evaluating Macie findings with AWS Security Hub.

Amazon GuardDuty

Amazon GuardDuty is a security monitoring service that analyzes and processes certain types of AWS logs, such as AWS CloudTrail data event logs for Amazon S3 and CloudTrail management event logs. It uses threat intelligence feeds, such as lists of malicious IP addresses and domains, and machine learning to identify unexpected and potentially unauthorized and malicious activity within your AWS environment.

To learn more about GuardDuty, see the Amazon GuardDuty User Guide.

To learn about additional AWS security services, see Security, Identity, and Compliance on AWS.

Getting started with Macie

This tutorial provides an introduction to Amazon Macie. You'll learn how to enable Macie for your AWS account. You'll also learn how to assess your Amazon Simple Storage Service (Amazon S3) security posture and configure key settings and resources for discovering and reporting sensitive data in your S3 buckets.

Tasks

Before you begin

When you sign up for Amazon Web Services (AWS), your account is automatically signed up for all AWS services, including Amazon Macie. However, to enable and use Macie, you first have to set up permissions that allow you to access the Amazon Macie console and API operations. You or your AWS administrator can do this by using AWS Identity and Access Management (IAM) to attach the AWS managed policy named `AmazonMacieFullAccess` to your IAM identity. To learn more, see AWS managed policies for Macie.

Step 1: Enable Macie

After you set up the required permissions, you can enable Amazon Macie for your AWS account. Follow these steps to enable Macie for your account.

To enable Macie

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to enable and use Macie.
3. On the Amazon Macie page, choose **Get started**.

4. (Optional) When you enable Macie, Macie automatically creates a service-linked role that allows it to call other AWS services and monitor AWS resources on your behalf. To review the permissions policy for this role, choose **View role permissions** on the console. To learn more about this role, see Using service-linked roles for Macie.
5. Choose **Enable Macie**.

Within minutes, Macie automatically generates and begins maintaining an inventory of your S3 general purpose buckets in the current Region. Macie also begins evaluating and monitoring the buckets for security and access control. To learn more, see Monitoring data security and privacy.

Depending on your account settings, Macie also begins performing automated sensitive data discovery for your S3 buckets. Macie begins to continually identify, select, and analyze representative objects in your buckets, inspecting the objects for sensitive data. As the analyses progress, Macie provides statistics and other results that you can review, typically within 48 hours. You can customize the analyses. To learn more, see Performing automated sensitive data discovery.

To review aggregated statistics for your Amazon S3 data, choose **Summary** in the navigation pane on the console. To review details about individual S3 buckets in your inventory, choose **S3 buckets** in the navigation pane. To then display a bucket's details, choose the bucket. The details panel displays statistics and other information that provide insight into the security, privacy, and sensitivity of the bucket's data. To learn about these details, see Reviewing your S3 bucket inventory.

Step 2: Configure a repository for sensitive data discovery results

With Amazon Macie, you can discover sensitive data in S3 buckets in two ways: by configuring Macie to perform automated sensitive data discovery and by running sensitive data discovery jobs. A *sensitive data discovery job* is a job that you create to analyze objects in S3 buckets to determine whether the objects contain sensitive data.

Macie creates a record for each S3 object that it analyzes when you run sensitive data discovery jobs or it performs automated sensitive data discovery. These records, referred to as *sensitive data discovery results*, log details about the analysis of individual objects. Macie also creates sensitive data discovery results for objects that it can't analyze due to errors or issues. Sensitive data discovery results provide you with analysis records that can be helpful for data privacy and protection audits or investigations.

Macie stores your sensitive data discovery results for only 90 days. To access the results and enable long-term storage and retention of them, configure Macie to store the results in an S3 bucket. You should do this within 30 days of enabling Macie. After you do this, the bucket can serve as a definitive, long-term repository for all of your sensitive data discovery results.

To learn how to configure this repository, see Storing and retaining sensitive data discovery results.

Step 3: Explore sample findings

In Amazon Macie, there are two categories of findings, *policy findings* and *sensitive data findings*. Macie creates a policy finding when the policies or settings for an S3 general purpose bucket are changed in a way that reduces the security or privacy of the bucket and the bucket's objects. Macie creates a sensitive data finding when it detects sensitive data in an S3 object. Within each category, there are multiple types of findings.

To explore and learn about the different categories and types of findings that Macie provides, optionally create and review sample findings. Sample findings use example data and placeholder values to demonstrate the kinds of information that Macie might include in each type of finding.

Follow these steps to create and review sample findings.

To create and review sample findings

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Settings**.
3. Under **Sample findings**, choose **Generate sample findings**. Macie generates one sample finding for each type of finding that Macie supports.
4. In the navigation pane, choose **Findings**. The **Findings** page displays findings for your account in the current AWS Region. This includes the sample findings that you created in the preceding step.
5. On the **Findings** page, locate findings whose type begins with **[SAMPLE]**.
6. To review the details of a particular sample finding, choose the finding. The details panel displays the finding's details.

To learn about each type of finding, see Types of findings. To learn more about creating and reviewing sample findings, see Working with sample findings.

Step 4: Create a job to discover sensitive data

To discover and report sensitive data in S3 buckets, you can run sensitive data discovery jobs. A *sensitive data discovery job* is a job that you create to analyze objects in S3 buckets to determine whether the objects contain sensitive data. Unlike automated sensitive data discovery, you define the breadth and depth of the analysis. You also specify how often to run a job—once or periodically on a scheduled basis.

Follow these steps to create a job that runs once, immediately after you create it, and uses default settings. To learn how to create a job that runs periodically or uses custom settings, see Creating a sensitive data discovery job.

To create a sensitive data discovery job

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Jobs**.
3. Choose **Create job**.
4. For the **Choose S3 buckets** step, choose **Select specific buckets**. Then, in the table, select the checkbox for each S3 bucket that you want the job to analyze.

   The table provides an inventory of your S3 general purpose buckets in the current AWS Region. To find specific buckets more easily, enter filter criteria in the filter box above the table. You can also sort the table by choosing a column heading.
5. When you finish selecting buckets, choose **Next**.
6. For the **Review S3 buckets** step, review and verify your bucket selections, and then choose **Next**.
7. For the **Refine the scope** step, choose **One-time job**, and then choose **Next**.
8. For the **Select managed data identifiers** step, choose **Recommended**. Optionally review the table of managed data identifiers that we recommend for jobs, and then choose **Next**.

   A *managed data identifier* is a set of built-in criteria and techniques that are designed to detect a specific type of sensitive data—for example, credit card numbers, AWS secret access keys, or passport numbers for a particular country or region. To learn more, see Using managed data identifiers.
9. For the **Select custom data identifiers** step, choose **Next**.

A *custom data identifier* is a set of criteria that you define to detect sensitive data—a regular expression (*regex*) that defines a text pattern to match and, optionally, character sequences and a proximity rule that refine the results. To learn more, see Building custom data identifiers.

10. For the **Select allow lists** step, choose **Next**.

    In Macie, an *allow list* specifies text or a text pattern that you want Macie to ignore when it inspects S3 objects for sensitive data. These are typically sensitive data exceptions for particular scenarios or environments. To learn more, see Defining sensitive data exceptions with allow lists.

11. For the **Enter general settings** step, enter a name and, optionally, a description of the job. Then choose **Next**.

12. For the **Review and create** step, review the job's configuration settings and verify that they're correct.

    You can also review the total estimated cost (in US dollars) of running the job. The estimate can help you determine whether to adjust the job's settings before you save the job. To learn more, see Forecasting the cost of a sensitive data discovery job.

13. When you finish reviewing and verifying the job's settings, choose **Submit**.

Macie immediately starts running the job. To learn how to monitor the job, see checking the status of sensitive data discovery jobs.

Step 5: Review findings

Amazon Macie automatically monitors your S3 general purpose buckets for security and access control, and it creates policy findings to report potential issues with the security or privacy of the buckets. If you run a sensitive data discovery job or configure Macie to perform automated sensitive data discovery, Macie creates sensitive data findings to report sensitive data that it detects in S3 objects.

Follow these steps to review findings.

To review findings

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.

2. In the navigation pane, choose **Findings**. The **Findings** page displays findings for your account in the current AWS Region.
3. To filter the findings by specific criteria, enter the criteria in the filter box above the table.
4. To review the details of a particular finding, choose the finding. The details panel displays the finding's details.

To learn more about findings, including how to group and filter them, see Reviewing and analyzing findings.

Concepts and terminology in Macie

In Amazon Macie, we build on common AWS concepts and terminology and use these additional terms.

account

A standard AWS account that contains your AWS resources and the identities that can access those resources.

To use Macie, you sign in to AWS with your AWS account credentials, select the AWS Region in which you want to use Macie, and then enable Macie for your AWS account in that Region. For more information, see Getting started with Macie.

There are three types of accounts in Macie:

- **Administrator account** – This type of account manages Macie accounts for an organization. An *organization* is a set of Macie accounts that are associated with each other and centrally managed as a group of related accounts in a specific AWS Region.
- **Member account** – This type of account is associated with and managed by the Macie administrator account for an organization.
- **Standalone account** – This type of account is neither an administrator nor a member account. It isn't part of an organization.

You can add Macie accounts to an organization in two ways: by integrating Macie with AWS Organizations or by sending and accepting Macie membership invitations. For more information, see Managing multiple accounts.

administrator account

In Macie, an account that manages Macie accounts for an organization. An *organization* is a set of Macie accounts that are associated with each other and centrally managed as a group of related accounts in a specific AWS Region.

Users of a Macie administrator account have access to Amazon Simple Storage Service (Amazon S3) inventory data, policy findings, and certain Macie settings and resources for all the accounts in their organization. They can also perform automated sensitive data discovery and run sensitive data

discovery jobs to detect sensitive data in S3 buckets that the accounts own. Depending on how an account is designated as an administrator account, they may also be able to perform additional tasks for other accounts in their organization.

For more information, see Managing multiple accounts.

allow list

In Macie, an allow list specifies text or a text pattern that you want Macie to ignore when it inspects S3 objects for sensitive data.

You can create two types of allow lists in Macie: a plaintext file that lists specific words and other kinds of character sequences to ignore, or a regular expression (*regex*) that defines a text pattern to ignore. If an object contains text that matches an entry or pattern in an allow list, Macie doesn't report the text in sensitive data findings, statistics, and other types of results. This is the case even if the text matches the criteria of a managed data identifier or a custom data identifier.

For more information, see Defining sensitive data exceptions with allow lists.

automated sensitive data discovery

A series of automated analysis activities that Macie continually performs to identify and select representative objects from S3 buckets, and inspect the selected objects for sensitive data.

As the analyses progress, Macie produces records of the sensitive data that it finds (sensitive data findings) and the analysis that it performs (sensitive data discovery results). Macie also updates statistics and other information that it provides about Amazon S3 data.

For more information, see Performing automated sensitive data discovery.

AWS Security Finding Format (ASFF)

A standardized JSON format for the contents of findings that are published to or generated by AWS Security Hub. The ASFF includes details about the source of a security issue, the affected resources, and the status of a finding.

For information about ASFF, see AWS Security Finding Format (ASFF) in the *AWS Security Hub User Guide*. For information about publishing Macie findings to Security Hub, see Evaluating findings with AWS Security Hub.

classifiable bytes or size

In the S3 bucket statistics that Macie provides, the total storage size of all the classifiable objects in an S3 bucket.

If versioning is enabled for a bucket, this value is based on the storage size of the latest version of each classifiable object in the bucket. If an object is a compressed file, this value doesn't reflect the actual size of the file's contents after the file is decompressed.

For more information, see Reviewing your S3 bucket inventory and Assessing your Amazon S3 security posture.

classifiable object

An S3 object that Macie can analyze to detect sensitive data.

When calculating S3 bucket statistics, Macie determines that an object is *classifiable* based on the object's storage class and file name extension. An object is *classifiable* if it uses a supported Amazon S3 storage class and has a file name extension for a supported file or storage format.

For more information, see Reviewing your S3 bucket inventory and Supported storage classes and formats.

For sensitive data discovery, Macie determines that an object is *classifiable* based on the object's storage class, file name extension, and contents. An object is *classifiable* if: it uses a supported Amazon S3 storage class, it has a file name extension for a supported file or storage format, and Macie verified that it can extract and analyze data from the object.

For more information, see Discovering sensitive data and Supported storage classes and formats.

custom data identifier

A set of criteria that you define to detect sensitive data.

The criteria consist of a regular expression (*regex*) that defines a text pattern to match and, optionally, character sequences and a proximity rule that refine the results. The character sequences can be:

- *Keywords*, which are words or phrases that must be in proximity of text that matches the regex, or

- *Ignore words*, which are words or phrases to exclude from the results.

In addition to detection criteria, you can define custom severity settings for the sensitive data findings that a custom data identifier produces.

For more information, see Building custom data identifiers.

filter rule

A set of attribute-based filter criteria that you create and save to analyze findings on the Amazon Macie console. Filter rules can help you perform consistent analysis of findings that have specific characteristics, such as all high-severity findings that report a specific type of sensitive data.

For more information, see Defining filter rules.

finding

A detailed report of sensitive data that Macie found in an S3 object or a potential issue with the security or privacy of an S3 general purpose bucket. Each finding provides details such as a severity rating, information about the affected resource, and when Macie found the data or issue.

Macie generates two categories of findings: sensitive data findings, for sensitive data that Macie detects in S3 objects, and policy findings, for potential issues that Macie detects with the security and access control settings for S3 buckets. Within each category, there are specific types of findings.

For more information, see Types of findings.

finding event

An Amazon EventBridge event that contains the details of a sensitive data finding or policy finding.

Macie automatically publishes sensitive data findings and policy findings to Amazon EventBridge as *events*. An event is a JSON object that conforms to the EventBridge schema for AWS events. You can use these events to monitor, process, and act upon findings by using other applications, services, and systems.

For more information, see Processing findings with Amazon EventBridge and Amazon EventBridge event schema for findings.

job

See sensitive data discovery job.

managed data identifier

A set of built-in criteria and techniques that are designed to detect a specific type of sensitive data. Examples of sensitive data include credit card numbers, AWS secret access keys, or passport numbers for a particular country or region. These identifiers can detect a large and growing list of sensitive data types for many countries and regions.

For more information, see Using managed data identifiers.

member account

A Macie account that's managed by the designated Macie administrator account for an organization. An *organization* is a set of Macie accounts that are associated with each other and centrally managed as a group of related accounts in a specific AWS Region.

An account can become a member account in two ways: by integrating Macie with the account's organization in AWS Organizations or by accepting a Macie membership invitation.

If you have a member account, your Macie administrator has access to Amazon S3 inventory data, policy findings, and certain Macie settings and resources for your account. Your administrator can also perform automated sensitive data discovery and run sensitive data discovery jobs to detect sensitive data in your S3 buckets. They may also be able to perform additional tasks for your account, depending on how your account became a member account.

For more information, see Managing multiple accounts.

organization

A set of Macie accounts that are associated with each other and centrally managed as a group of related accounts in a specific AWS Region.

Each organization consists of a designated Macie administrator account and one or more associated member accounts. The administrator account can access certain Macie settings, data, and resources for member accounts. You can create an organization in two ways: by integrating Macie with AWS Organizations or by sending and accepting membership invitations in Macie.

For more information, see Managing multiple accounts.

policy finding

A detailed report of a potential policy violation or issue with the security and access control settings for an S3 general purpose bucket. The details include a severity rating, information about the affected resource, and when Macie found the issue.

Macie generates policy findings when the policies or settings for an S3 general purpose bucket are changed in a way that reduces the security or privacy of the bucket and the bucket's objects. Macie generates these findings as part of its ongoing monitoring activities for your Amazon S3 data. Macie can generate several types of policy findings.

For more information, see Types of findings and Monitoring data security and privacy.

sample finding

A finding that uses example data and placeholder values to demonstrate the kinds of information that a finding might contain.

For more information, see Working with sample findings.

sensitive data finding

A detailed report of sensitive data that Macie found in an S3 object. The details include a severity rating, information about the affected resource, the type and number of occurrences of the sensitive data that Macie found, and when Macie found the sensitive data.

Macie generates sensitive data findings if it detects sensitive data in S3 objects that it analyzes when you run sensitive data discovery jobs or it performs automated sensitive data discovery. Macie can generate several types of sensitive data findings.

For more information, see Types of findings and Discovering sensitive data.

sensitive data discovery job

Also referred to as a *job*, a series of automated processing and analysis tasks that Macie performs to detect and report sensitive data in S3 objects. When you create a job, you specify how often you want the job to run, and you define the scope and nature of the job's analysis.

When a job runs, Macie produces records of the sensitive data that it finds (sensitive data findings) and the analysis that it performs (sensitive data discovery results). Macie also publishes logging data to Amazon CloudWatch Logs.

For more information, see Running sensitive data discovery jobs.

sensitive data discovery result

A record that logs details about the analysis that Macie performed on an S3 object to determine whether the object contains sensitive data. Macie generates and writes these records to JSON Lines (.jsonl) files, which it encrypts and stores in an S3 bucket that you specify. The records adhere to a standardized schema.

When you run a sensitive data discovery job or Macie performs automated sensitive data discovery, Macie creates a sensitive data discovery result for each object that's included in the scope of the analysis. This includes:

- Objects that Macie finds sensitive data in, and therefore also produce sensitive data findings.
- Objects that Macie doesn't find sensitive data in, and therefore don't produce sensitive data findings.
- Objects that Macie can't analyze due to errors or issues such as permissions settings or use of an unsupported file or storage format.

For more information, see Storing and retaining sensitive data discovery results.

session

A resource that represents the Macie service for a specific AWS account in a specific AWS Region. An AWS account can have only one Macie session in each Region.

When you enable Macie for the first time, the service generates a Macie session for your account in the current Region. It also assigns a unique identifier to that session. The session enables Macie to become operational for your account in the Region.

standalone account

A Macie account that's neither an administrator nor a member account in an organization. The account isn't part of an organization.

suppressed finding

A finding that was archived automatically by a suppression rule. That is to say, Macie automatically changed the status of the finding to *archived* because the finding matched the criteria of a suppression rule when Macie generated the finding.

For more information, see Suppressing findings.

suppression rule

A set of attribute-based filter criteria that you create and save to archive (*suppress*) findings automatically. Suppression rules are helpful in situations where you've reviewed a class of findings and don't want to be notified of them again.

If you suppress findings with a suppression rule, Macie continues to generate findings that match the rule's criteria. However, Macie automatically changes the status of the findings to *archived*. This means that the findings don't appear by default on the Amazon Macie console and Macie doesn't publish them to other AWS services.

For more information, see Suppressing findings.

unclassifiable bytes or size

In the S3 bucket statistics that Macie provides, the total storage size of all the unclassifiable objects in an S3 bucket.

If versioning is enabled for a bucket, this value is based on the storage size of the latest version of each unclassifiable object in the bucket. If an object is a compressed file, this value doesn't reflect the actual size of the file's contents after the file is decompressed.

For more information, see Reviewing your S3 bucket inventory and Assessing your Amazon S3 security posture.

unclassifiable object

An S3 object that Macie can't analyze to detect sensitive data.

When calculating S3 bucket statistics, Macie determines that an object is *unclassifiable* based on the object's storage class and file name extension. An object is *unclassifiable* if it doesn't use a

supported Amazon S3 storage class or doesn't have a file name extension for a supported file or storage format.

For more information, see Reviewing your S3 bucket inventory and Supported storage classes and formats.

For sensitive data discovery, Macie determines that an object is *unclassifiable* based on the object's storage class, file name extension, and contents. An object is *unclassifiable* if: it doesn't use a supported Amazon S3 storage class, it doesn't have a file name extension for a supported file or storage format, or Macie wasn't able to extract and analyze data from the object. For example, the object is a malformed file.

For more information, see Discovering sensitive data and Supported storage classes and formats.

Monitoring data security and privacy with Macie

When you enable Amazon Macie for your AWS account, Macie automatically generates and begins maintaining an inventory of your Amazon Simple Storage Service (Amazon S3) general purpose buckets in the current AWS Region. Macie also begins evaluating and monitoring the buckets for security and access control. If Macie detects an event that reduces the security or privacy of a bucket, Macie creates a policy finding for you to review and remediate as necessary.

To also evaluate and monitor the S3 buckets for the presence of sensitive data, you can create and run sensitive data discovery jobs. Sensitive data discovery jobs can perform incremental analysis of bucket objects on a daily, weekly, or monthly basis. If Macie detects sensitive data in an S3 object, Macie creates a sensitive data finding to notify you of the sensitive data that it found. Depending on your account settings, you can also configure Macie to perform automated sensitive data discovery. Automated sensitive data discovery uses sampling techniques to continually identify, select, and analyze representative objects in your buckets. For more information about both options, see Discovering sensitive data.

Macie also provides constant visibility into the security and privacy of your Amazon S3 data. To assess the security posture of your data and determine where to take action, you can use the **Summary** dashboard on the console. The dashboard provides a snapshot of aggregated statistics for your Amazon S3 data. The statistics include data for key security metrics such as the number of general purpose buckets that are publicly accessible or shared with other AWS accounts. The dashboard also displays groups of aggregated findings data for your account—for example, the names of 1–5 buckets that have the most findings for the preceding seven days. You can drill down on each statistic to review its supporting data. To query the statistics programmatically, use the GetBucketStatistics operation of the Amazon Macie API.

For deeper analysis and evaluation, Macie provides detailed information and statistics for individual S3 buckets in your inventory. This includes breakdowns of each bucket's public access and encryption settings, and the size and number of objects that Macie can analyze to detect sensitive data in the bucket. The inventory also indicates whether you configured sensitive data discovery jobs or automated sensitive data discovery to analyze objects in a bucket. If you have, it indicates when that analysis most recently occurred. You can browse, sort, and filter the inventory by using the Amazon Macie console or the DescribeBuckets operation of the Amazon Macie API.

If you're the Macie administrator for an organization, you can access statistical and other data about S3 buckets that your member accounts own. You can also access policy findings that Macie

generates for the buckets, and inspect the buckets for sensitive data. This means that you can use Macie to assess and monitor the overall security posture of your organization's Amazon S3 data estate. For more information, see Managing multiple accounts.

Topics

How Macie monitors Amazon S3 data security

When you enable Amazon Macie for your AWS account, Macie creates an AWS Identity and Access Management (IAM) service-linked role for your account in the current AWS Region. The permissions policy for this role allows Macie to call other AWS services and monitor AWS resources on your behalf. By using this role, Macie generates and maintains an inventory of your Amazon Simple Storage Service (Amazon S3) general purpose buckets in the Region. Macie also monitors and evaluates the buckets for security and access control.

If you're the Macie administrator for an organization, the inventory includes statistical and other data about S3 buckets for your account and member accounts in your organization. With this data, you can use Macie to monitor and evaluate your organization's security posture across your Amazon S3 data estate. For more information, see Managing multiple accounts.

Topics

Key components

Amazon Macie uses a combination of features and techniques to provide and maintain inventory data for your S3 general purpose buckets, and to monitor and evaluate the buckets for security and access control.

Gathering metadata and calculating statistics

To generate and maintain metadata and statistics for your bucket inventory, Macie retrieves bucket and object metadata directly from Amazon S3. For each bucket, the metadata includes:

- General information about the bucket, such as the bucket's name, Amazon Resource Name (ARN), creation date, encryption settings, tags, and the account ID for the AWS account that owns the bucket.
- Account-level permissions settings that apply to the bucket, such as the block public access settings for the account.
- Bucket-level permissions settings for the bucket, such as the block public access settings for the bucket and settings that derive from a bucket policy or access control list (ACL).
- Shared access and replication settings for the bucket, including whether bucket data is replicated to or shared with AWS accounts that aren't part of your organization.
- Object counts and settings for objects in the bucket, such as the number of objects in the bucket and breakdowns of object counts by encryption type, file type, and storage class.

Macie provides this information to you directly. Macie also uses the information to calculate statistics and provide assessments of the security and privacy of your bucket inventory overall and individual buckets in your inventory. For example, you can find the total storage size and number of buckets in your inventory, the total storage size and number of objects in those buckets, and the total storage size and number of objects that Macie can analyze to detect sensitive data in the buckets.

By default, metadata and statistics include data for any object parts that exist due to incomplete multipart uploads. If you manually refresh object metadata for a specific bucket, Macie recalculates statistics for the bucket and your bucket inventory overall, and excludes data for object parts from the recalculated values. The next time Macie retrieves bucket and object metadata from Amazon S3 as part of the daily refresh cycle, Macie updates your inventory data and includes data for the object parts again. For information about when Macie retrieves bucket and object metadata, see Data refreshes.

It's important to note that Macie can't analyze object parts to detect sensitive data. Amazon S3 must first finish assembling the parts into one or more objects for Macie to analyze. For information about multipart uploads and object parts, including how to delete parts automatically with lifecycle rules, see Uploading and copying objects using multipart upload in the *Amazon Simple Storage Service User Guide*. To identify buckets that contain object parts, you can refer to *incomplete multipart upload* metrics in Amazon S3 Storage Lens. For more

information, see Assessing your storage activity and usage in the *Amazon Simple Storage Service User Guide*.

Monitoring bucket security and privacy

To help ensure the accuracy of bucket-level data in your inventory, Macie monitors and analyzes certain AWS CloudTrail events that can occur for Amazon S3 data. If a relevant event occurs, Macie updates the appropriate inventory data.

For example, if you enable block public access settings for a bucket, Macie updates all data about the bucket's public access settings. Similarly, if you add or update the bucket policy for a bucket, Macie analyzes the policy and updates the appropriate data in your inventory.

If Macie determines that an event reduces the security or privacy of a bucket, Macie also creates a policy finding for you to review and remediate as necessary.

Macie monitors and analyzes data for the following CloudTrail events:

- **Account-level events** – DeletePublicAccessBlock and PutPublicAccessBlock
- **Bucket-level events** – CreateBucket, DeleteAccountPublicAccessBlock, DeleteBucket, DeleteBucketEncryption, DeleteBucketPolicy, DeleteBucketPublicAccessBlock, DeleteBucketReplication, DeleteBucketTagging, PutAccountPublicAccessBlock, PutBucketAcl, PutBucketEncryption, PutBucketPolicy, PutBucketPublicAccessBlock, PutBucketReplication, PutBucketTagging, and PutBucketVersioning

You can't enable monitoring for additional CloudTrail events or disable monitoring for any of the preceding events. For detailed information about corresponding operations for the preceding events, see the Amazon Simple Storage Service API Reference.

> **Tip**
>
> To monitor object-level events, we recommend that you use the Amazon S3 protection feature of Amazon GuardDuty. This feature monitors object-level, Amazon S3 data events and analyzes them for malicious and suspicious activity. For more information, see GuardDuty S3 Protection in the *Amazon GuardDuty User Guide*.

Evaluating bucket security and access control

To evaluate bucket-level security and access control, Macie uses automated, logic-based reasoning to analyze resource-based policies that apply to a bucket. Macie also analyzes the

account- and bucket-level permissions settings that apply to a bucket. This analysis factors bucket policies, bucket-level ACLs, and block public access settings for the account and the bucket.

For resource-based policies, Macie uses Zelkova. Zelkova is an automated reasoning engine that translates AWS Identity and Access Management (IAM) policies into logical statements and runs a suite of general-purpose and specialized logical solvers (*satisfiability modulo theories*) against the decision problem. To learn more about the nature of the solvers that Zelkova uses, see Satisfiability Modulo Theories.

Macie applies Zelkova repeatedly to a resource-based policy, using increasingly specific queries to characterize the classes of behaviors that the policy allows. The analysis is designed to identify potential security risks for your Amazon S3 data and minimize false negatives. It doesn't include AWS Organizations authorization policies that define the maximum available permissions for your organization's resources, such as service control policies (SCPs) or resource control policies (RCPs). It also doesn't include key policies for associated AWS KMS keys. For example, if a bucket policy uses the s3:x-amz-server-side-encryption-aws-kms-key-id condition key to restrict write access to the bucket, Macie doesn't analyze the key policy for the specified key. This means that Macie might report that the bucket is publicly accessible, depending on other components of the bucket policy and Amazon S3 permissions settings that apply to the bucket.

In addition, when Macie assesses the security and privacy of a bucket, it doesn't examine access logs or analyze users, roles, and other relevant configurations for accounts. Instead, Macie analyzes and reports data for key settings that indicate *potential* security risks. For example, if a policy finding indicates that a bucket is publicly accessible, it doesn't necessarily mean that an external entity accessed the bucket. Similarly, if a policy finding indicates that a bucket is shared with an AWS account outside your organization, Macie doesn't attempt to determine whether this access is intended and safe. Instead, these findings indicate that an external entity can potentially access the bucket's data, which may be an unintended security risk.

If Macie reports that an external entity can potentially access an S3 bucket, we recommend that you review the bucket's policy and settings to determine whether this access is intended and safe. If applicable, also review policies and settings for associated resources, such as AWS KMS keys, and AWS Organizations authorization policies for your organization.

Important

To perform the preceding tasks for a bucket, the bucket must be an S3 general purpose bucket. Macie doesn't monitor or analyze S3 directory buckets.
In addition, Macie must be allowed to access the bucket. If a bucket's permissions settings prevent Macie from retrieving metadata for the bucket or the bucket's objects, Macie can only provide a subset of information about the bucket, such as the bucket's name and creation date. Macie can't perform any additional tasks for the bucket. For more information, see Allowing Macie to access S3 buckets and objects.
Macie can perform the preceding tasks for up to 10,000 buckets for an account. If you store more than 10,000 buckets in Amazon S3, Macie performs these tasks only for the 10,000 buckets that were most recently created or changed. For all other buckets, Macie doesn't maintain complete inventory data, evaluate or monitor the security and privacy of the buckets' data, or generate policy findings. Instead, Macie only provides a subset of information about the buckets.

Data refreshes

When you enable Amazon Macie for your AWS account, Macie retrieves metadata for your S3 general purpose buckets and objects directly from Amazon S3. Thereafter, Macie automatically retrieves bucket and object metadata directly from Amazon S3 on a daily basis as part of a daily refresh cycle.

Macie also retrieves bucket metadata directly from Amazon S3 when any of the following occurs:

- Macie detects a relevant AWS CloudTrail event.
- You refresh your inventory data by choosing refresh
  ( )
  on the Amazon Macie console. Depending on the size of your data estate, you can refresh the data as frequently as every five minutes.
- You submit a DescribeBuckets request to the Amazon Macie API programmatically and Macie has finished processing any preceding **DescribeBuckets** requests.

Macie can also retrieve the latest object metadata for a specific bucket if you choose to manually refresh that data. This can be helpful if you recently created

a bucket or made significant changes to a bucket's objects during the past 24 hours. To manually refresh object metadata for a bucket, choose refresh () in the **Object statistics** section of the bucket details panel on the **S3 buckets** page of the console. This feature is available for buckets that store 30,000 or fewer objects.

To determine when Macie most recently retrieved bucket or object metadata for your account, you can refer to the **Last updated** field on the console. This field appears on the **Summary** dashboard, on the **S3 buckets** page, and in the bucket details panel on the **S3 buckets** page. If you use the Amazon Macie API to query inventory data, the `lastUpdated` field provides this information. If you're the Macie administrator for an organization, the field indicates the earliest date and time when Macie retrieved the data for an account in your organization.

Each time Macie retrieves bucket or object metadata, Macie automatically updates the appropriate data in your inventory. If Macie detects differences that affect the security or privacy of a bucket, Macie immediately begins evaluating and analyzing the changes. When the analysis is complete, Macie updates the appropriate data in your inventory. If any differences reduce the security or privacy of a bucket, Macie also creates the appropriate policy findings for you to review and remediate as necessary. Macie does this for as many as 10,000 buckets for your account. If you have more than 10,000 buckets, Macie does this for the 10,000 buckets that were most recently created or changed. If you're the Macie administrator for an organization, this quota applies to each account in your organization, not your organization overall.

On rare occasions under certain conditions, latency and other issues might prevent Macie from retrieving bucket and object metadata. They might also delay notifications that Macie receives about changes to your bucket inventory or the permissions settings and policies for individual buckets. For example, delivery issues with CloudTrail events might cause delays. If this happens, Macie analyzes new and updated data the next time it performs the daily refresh, which is within 24 hours.

Considerations

As you use Amazon Macie to monitor and assess the security posture of your Amazon S3 data, keep the following in mind:

- Inventory data applies only to S3 general purpose buckets in the current AWS Region. To access the data for additional Regions, enable and use Macie in each additional Region.

- If you're the Macie administrator for an organization, you can access inventory data for a member account only if Macie is enabled for that account in the current Region.
- Macie can provide complete inventory data for no more than 10,000 buckets for an account. In addition, Macie can evaluate and monitor the security and privacy of no more than 10,000 buckets for an account. If your account exceeds this quota, Macie evaluates, monitors, and provides detailed information about the 10,000 buckets that were most recently created or changed. For all other buckets, Macie only provides a subset of information about the buckets.

  If your account approaches this quota, we notify you by creating an AWS Health event for your account. We also send email to the address that's associated with your account. We notify you again if your account exceeds the quota. If you're a Macie administrator, this quota applies to each account in your organization, not your organization overall.
- If a bucket's permissions settings prevent Macie from retrieving information about the bucket or the bucket's objects, Macie can't evaluate and monitor the security and privacy of the bucket's data or provide detailed information about the bucket. To help you identify a bucket where this is the case, Macie does the following:
  - In your bucket inventory on the console, Macie displays a warning icon () for the bucket.
  - For the bucket's details, Macie provides data for only a subset of fields: the account ID for the AWS account that owns the bucket; the bucket's name, Amazon Resource Name (ARN), creation date, and Region; and, the date and time when Macie most recently retrieved both bucket and object metadata for the bucket as part of the daily refresh cycle. If you query inventory data programmatically with the Amazon Macie API, Macie also provides an error code and message for the bucket.
  - In the **Summary** dashboard on the console, the bucket has a value of **Unknown** for **Public access**, **Encryption**, and **Sharing** statistics. In addition, Macie excludes the bucket when it calculates data for **Storage** and **Objects** statistics.
  - If you query aggregated statistics programmatically by using the GetBucketStatistics operation, the bucket has a value of unknown for many statistics and Macie excludes the bucket when it calculates object counts and storage size values.

  To investigate the issue, review the bucket's policy and permissions settings in Amazon S3. For example, the bucket might have a restrictive bucket policy. For more information, see Allowing Macie to access S3 buckets and objects.

- Data about access and permissions is limited to account- and bucket-level settings. It doesn't reflect object-level settings that determine access to specific objects in a bucket. For example, if public access is enabled for a specific object in a bucket, Macie doesn't report that the bucket or the bucket's objects are publicly accessible.

  To monitor object-level operations and identify potential security risks, we recommend that you use the Amazon S3 protection feature of Amazon GuardDuty. This feature monitors object-level, Amazon S3 data events and analyzes them for malicious and suspicious activity. For more information, see GuardDuty S3 Protection in the *Amazon GuardDuty User Guide*.
- If you manually refresh object metadata for a specific bucket:
  - Macie temporarily reports *Unknown* for encryption statistics that apply to the objects. The next time Macie performs the daily data refresh (within 24 hours), Macie re-evaluates the encryption metadata for the objects and reports quantitative data for the statistics again.
  - Macie temporarily excludes data for any object parts that the bucket contains due to incomplete multipart uploads. The next time Macie performs the daily data refresh (within 24 hours), Macie recalculates counts and storage size values for the bucket's objects and includes data for the parts in those calculations.
- In certain cases, Macie might not be able to determine whether a bucket is publicly accessible or shared, or requires server-side encryption of new objects. For example, a quota or temporary issue might prevent Macie from retrieving and analyzing the requisite data. Or Macie might not be able to fully determine whether one or more policy statements grant access to an external entity. In these cases, Macie reports *Unknown* for the relevant statistics and fields in your bucket inventory. To investigate these cases, review the bucket's policy and permissions settings in Amazon S3.

Also note that Macie generates policy findings only if the security or privacy of a bucket is reduced after you enable Macie for your account. For example, if you disable block public access settings for a bucket after you enable Macie, Macie generates a **Policy:IAMUser/S3BlockPublicAccessDisabled** finding for the bucket. However, if block public access settings were disabled for a bucket when you enabled Macie and they continue to be disabled, Macie doesn't generate a **Policy:IAMUser/S3BlockPublicAccessDisabled** finding for the bucket.

Assessing your Amazon S3 security posture with Macie

To assess the overall security posture of your Amazon Simple Storage Service (Amazon S3) data and determine where to take action, you can use the **Summary** dashboard on the Amazon Macie console.

The **Summary** dashboard provides a snapshot of aggregated statistics for your Amazon S3 data in the current AWS Region. The statistics include data for key security metrics such as the number of general purpose buckets that are publicly accessible or shared with other AWS accounts. The dashboard also displays groups of aggregated findings data for your account—for example, the types of findings that had the highest number of occurrences during the preceding seven days. If you're the Macie administrator for an organization, the dashboard provides aggregated statistics and data for all the accounts in your organization. You can optionally filter the data by account.

To perform deeper analysis, you can drill down and review the supporting data for individual items on the dashboard. You can also review and analyze your S3 bucket inventory by using the Amazon Macie console, or query and analyze inventory data programmatically by using the DescribeBuckets operation of the Amazon Macie API.

Topics

Displaying the Summary dashboard

On the Amazon Macie console, the **Summary** dashboard provides a snapshot of aggregated statistics and findings data for your Amazon S3 data in the current AWS Region. If you prefer to query the statistics programmatically, you can use the GetBucketStatistics operation of the Amazon Macie API.

To display the Summary dashboard

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Summary**. Macie displays the **Summary** dashboard.

3. To determine when Macie most recently retrieved bucket or object metadata from Amazon S3 for your account, refer to the **Last updated** field at the top of the dashboard. For more information, see Data refreshes.
4. To drill down and review the supporting data for an item on the dashboard, choose the item.

If you're the Macie administrator for an organization, the dashboard displays aggregated statistics and data for your account and member accounts in your organization. To filter the dashboard and display data only for a particular account, enter the account's ID in the **Account** box above the dashboard.

Understanding components of the Summary dashboard

On the **Summary** dashboard, statistics and data are organized into several sections. At the top of the dashboard, you'll find aggregated statistics that indicate how much data you store in Amazon S3, and how much of that data Amazon Macie can analyze to detect sensitive data. You can also refer to the **Last updated** field to determine when Macie most recently retrieved bucket or object metadata from Amazon S3 for your account. Additional sections provide statistics and recent findings data that can help you assess the security, privacy, and sensitivity of your Amazon S3 data in the current AWS Region.

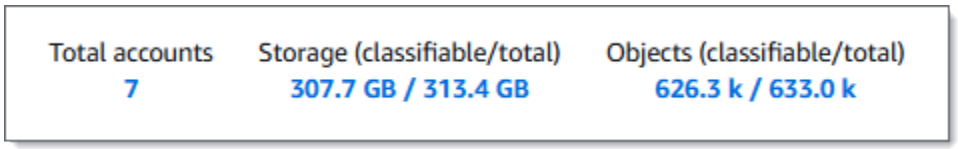
Statistics and data are organized into the following sections:

Storage and sensitive data discovery | Automated discovery and coverage issues | Data security | Top S3 buckets | Top finding types | Policy findings

As you review each section, optionally choose an item to drill down and review the supporting data. Also note that the dashboard doesn't include data for S3 directory buckets, only general purpose buckets. Macie doesn't monitor or analyze directory buckets.

Storage and sensitive data discovery

At the top of the dashboard, statistics indicate how much data you store in Amazon S3, and how much of that data Macie can analyze to detect sensitive data. The following image shows an example of these statistics for an organization with seven accounts.



Individual statistics in this section are:

- **Total accounts** – This field appears if you're the Macie administrator for an organization or you have a standalone Macie account. It indicates the total number of AWS accounts that own buckets in your bucket inventory. If you're a Macie administrator, this is the total number of Macie accounts that you manage for your organization. If you have a standalone Macie account, this value is *1*.

  Total S3 buckets – This field appears if you have a member account in an organization. It indicates the total number of general purpose buckets in your inventory, including buckets that don't store any objects.

- **Storage** – These statistics provide information about the storage size of objects in your bucket inventory:
  - **Classifiable** – The total storage size of all the objects that Macie can analyze in the buckets.
  - **Total** – The total storage size of all the objects in the buckets, including objects that Macie can't analyze.

  If any of the objects are compressed files, these values don't reflect the actual size of those files after they're decompressed. If versioning is enabled for any of the buckets, these values are based on the storage size of the latest version of each object in those buckets.

- **Objects** – These statistics provide information about the number of objects in your bucket inventory:
  - **Classifiable** – The total number of objects that Macie can analyze in the buckets.
  - **Total** – The total number of objects in the buckets, including objects that Macie can't analyze.

In the preceding statistics, data and objects are *classifiable* if they use a supported Amazon S3 storage class and they have a file name extension for a supported file or storage format. You can detect sensitive data in the objects by using Macie. For more information, see Supported storage classes and formats.

Note that **Storage** and **Objects** statistics don't include data about objects in buckets that Macie isn't allowed to access. For example, objects in buckets that have restrictive bucket policies. To identify buckets where this is the case, you can review your bucket inventory by using the **S3 buckets** table. If the warning icon (⚠) appears next to a bucket's name, Macie isn't allowed to access the bucket.

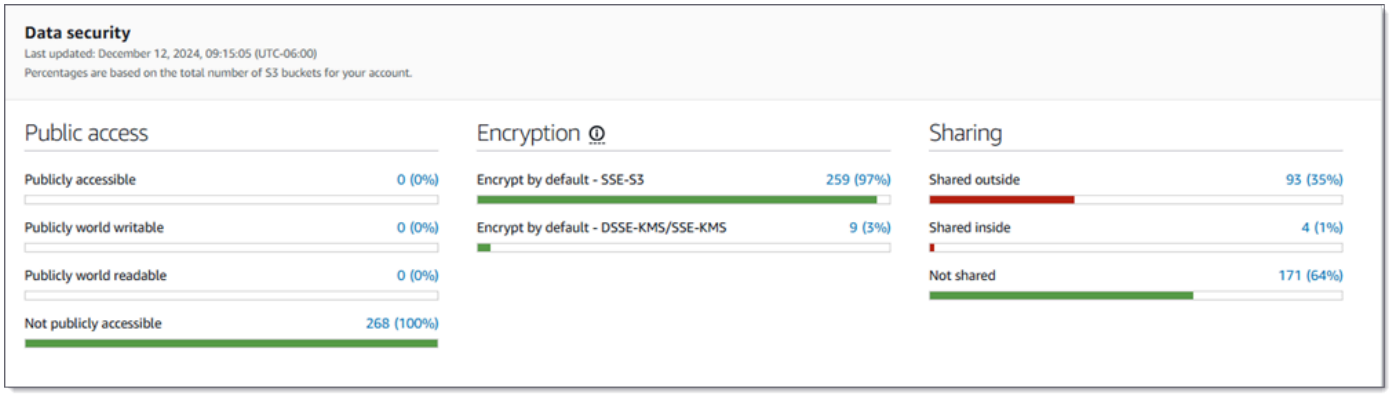
Automated discovery and coverage issues

If automated sensitive data discovery is enabled, these sections appear on the dashboard. They capture the status and results of automated sensitive data discovery activities that Macie has performed thus far for your Amazon S3 data. The following image shows an example of the statistics that these sections provide.

For details about these statistics, see Reviewing data sensitivity statistics on the Summary dashboard.

Data security

This section provides statistics that indicate potential security and privacy risks for your Amazon S3 data. The following image shows an example of the statistics in this section.

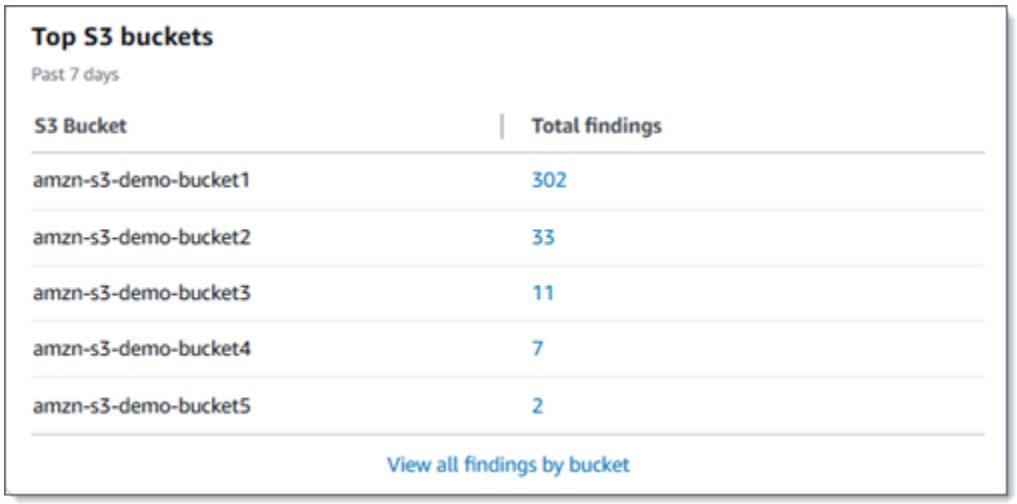


For details about these statistics, see Understanding data security statistics on the Summary dashboard.

Top S3 buckets

This section lists the S3 buckets that generated the most findings of any type during the preceding seven days, for as many as five buckets. It also indicates the number of findings that

Macie created for each bucket. The following image shows an example of the data that this section provides.

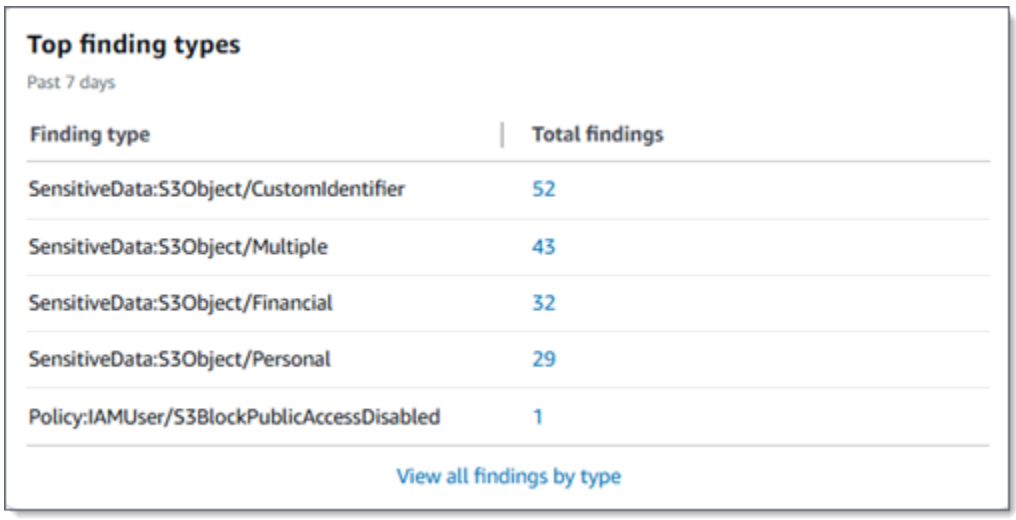


To display and optionally drill down on all the findings for a bucket for the preceding seven days, choose the value in the **Total findings** field. To display all current findings for all of your buckets, grouped by bucket, choose **View all findings by bucket**.

This section is empty if Macie didn't create any findings during the preceding seven days. Or all the findings that were created during the preceding seven days were suppressed by a suppression rule.

Top finding types

This section lists the types of findings that had the highest number of occurrences during the preceding seven days, for as many as five types of findings. It also indicates the number of findings that Macie created for each type. The following image shows an example of the data that this section provides.



To display and optionally drill down on all findings of a particular type for the preceding seven days, choose the value in the **Total findings** field. To display all current findings, grouped by finding type, choose **View all findings by type**.

This section is empty if Macie didn't create any findings during the preceding seven days. Or all the findings that were created during the preceding seven days were suppressed by a suppression rule.

Policy findings

This section lists the policy findings that Macie created or updated most recently, for as many as ten findings. The following image shows an example of the data that this section provides.

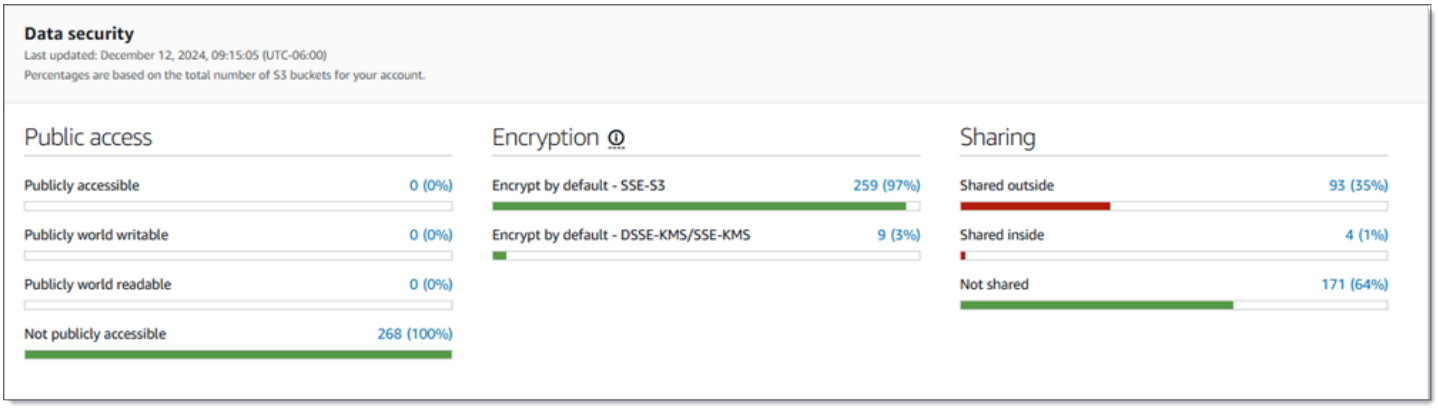
To display the details of a particular finding, choose the finding.

This section is empty if Macie didn't create or update any policy findings during the preceding seven days. Or all the policy findings that were created or updated during the preceding seven days were suppressed by a suppression rule.

Understanding data security statistics on the Summary dashboard

The **Data security** section of the **Summary** dashboard provides statistics that can help you identify and investigate potential security and privacy risks for your Amazon S3 data in the current AWS Region. For example, you can use this data to identify general purpose buckets that are publicly accessible or shared with other AWS accounts.

If automated sensitive data discovery is disabled, storage and sensitive data discovery statistics at the top of this section indicate how much data you store in Amazon S3, and how much of that data Amazon Macie can analyze to detect sensitive data. Additional statistics are organized into three areas, as shown in the following image.



As you review each area, optionally choose an item to drill down and review the supporting data. Also note that the statistics don't include data for S3 directory buckets, only general purpose buckets. Macie doesn't monitor or analyze directory buckets.

Individual statistics in each area are as follows.

Public access

These statistics indicate how many S3 buckets are or aren't publicly accessible:

- **Publicly accessible** – The number and percentage of buckets that allow the general public to have read or write access to the bucket.
- **Publicly world writable** – The number and percentage of buckets that allow the general public to have write access to the bucket.
- **Publicly world readable** – The number and percentage of buckets that allow the general public to have read access to the bucket.
- **Not publicly accessible** – The number and percentage of buckets that don't allow the general public to have read or write access to the bucket.

To calculate each percentage, Macie divides the number of applicable buckets by the total number of buckets in your bucket inventory.

To determine the values in this area, Macie analyzes a combination of account- and bucket-level settings for each bucket: the block public access settings for the account; the block public access settings for the bucket; the bucket policy for the bucket; and, the access control list (ACL) for the bucket. For information about these settings, see Access control and Blocking public access to your Amazon S3 storage in the *Amazon Simple Storage Service User Guide*.

In certain cases, the **Public access** area also displays values for **Unknown**. If these values appear, Macie wasn't able to evaluate the public access settings for the specified number and percentage of buckets. For example, a temporary issue or the buckets' permissions settings prevented Macie from retrieving the requisite data. Or Macie wasn't able to fully determine whether one or more policy statements allow an external entity to access the buckets. This can also be the case for buckets that exceed the quota for preventative control monitoring. Macie evaluates and monitors the security and privacy of no more than 10,000 buckets for an account—the 10,000 buckets that were most recently created or changed.

Encryption

These statistics indicate how many S3 buckets are configured to apply certain types of server-side encryption to objects that are added to the buckets:

- **Encrypt by default – SSE-S3** – The number and percentage of buckets whose default encryption settings are configured to encrypt new objects with an Amazon S3 managed key. For these buckets, new objects are encrypted automatically using SSE-S3 encryption.
- **Encrypt by default – DSSE-KMS/SSE-KMS** – The number and percentage of buckets whose default encryption settings are configured to encrypt new objects with an AWS KMS key, either an AWS managed key or a customer managed key. For these buckets, new objects are encrypted automatically using DSSE-KMS or SSE-KMS encryption.

To calculate each percentage, Macie divides the number of applicable buckets by the total number of buckets in your bucket inventory.

To determine the values in this area, Macie analyzes the default encryption settings for each bucket. Starting January 5, 2023, Amazon S3 automatically applies server-side encryption with Amazon S3 managed keys (SSE-S3) as the base level of encryption for objects that are added to buckets. You can optionally configure a bucket's default encryption settings to instead use server-side encryption with an AWS KMS key (SSE-KMS) or dual-layer server-side encryption

with an AWS KMS key (DSSE-KMS). For information about default encryption settings and options, see Setting default server-side encryption behavior for S3 buckets in the *Amazon Simple Storage Service User Guide*.

In certain cases, the **Encryption** area also displays values for **Unknown**. If these values appear, Macie wasn't able to evaluate the default encryption settings for the specified number and percentage of buckets. For example, a temporary issue or the buckets' permissions settings prevented Macie from retrieving the requisite data. Or the buckets exceed the quota for preventative control monitoring. Macie evaluates and monitors the security and privacy of no more than 10,000 buckets for an account—the 10,000 buckets that were most recently created or changed.

Sharing

These statistics indicate how many S3 buckets are or aren't shared with other AWS accounts, Amazon CloudFront origin access identities (OAIs), or CloudFront origin access controls (OACs):

- **Shared outside** – The number and percentage of buckets that are shared with one or more of the following or any combination of the following: a CloudFront OAI, a CloudFront OAC, or an account that isn't in the same organization.
- **Shared inside** – The number and percentage of buckets that are shared with one or more accounts in the same organization. These buckets aren't shared with CloudFront OAIs or OACs.
- **Not shared** – The number and percentage of buckets that aren't shared with other accounts, CloudFront OAIs, or CloudFront OACs.

To calculate each percentage, Macie divides the number of applicable buckets by the total number of buckets in your bucket inventory.

To determine whether buckets are shared with other AWS accounts, Macie analyzes the bucket policy and ACL for each bucket. In addition, an *organization* is defined as a set of Macie accounts that are centrally managed as a group of related accounts through AWS Organizations or by Macie invitation. For information about Amazon S3 options for sharing buckets, see Access control in the *Amazon Simple Storage Service User Guide*.

> **Note**
>
> In certain cases, Macie might incorrectly report that a bucket is shared with an AWS account that isn't in the same organization. This can occur if Macie isn't able to fully

evaluate the relationship between the `Principal` element in a bucket's policy and certain AWS global condition context keys or Amazon S3 condition keys in the `Condition` element of the policy. This can be the case for the following condition keys: `aws:PrincipalAccount`, `aws:PrincipalArn`, `aws:PrincipalOrgID`, `aws:PrincipalOrgPaths`, `aws:PrincipalTag`, `aws:PrincipalType`, `aws:SourceAccount`, `aws:SourceArn`, `aws:SourceIp`, `aws:SourceOrgID`, `aws:SourceOrgPaths`, `aws:SourceVpc`, `aws:SourceVpce`, `aws:userid`, `s3:DataAccessPointAccount`, and `s3:DataAccessPointArn`.
To determine whether this is the case for individual buckets, choose the **Shared outside** statistic on the dashboard. In the table that appears, note the name of each bucket. Then use Amazon S3 to review each bucket's policy and determine whether the shared access settings are intended and safe.

To determine whether buckets are shared with CloudFront OAIs or OACs, Macie analyzes the bucket policy for each bucket. A CloudFront OAI or OAC allows users to access a bucket's objects through one or more specified CloudFront distributions. For information about CloudFront OAIs and OACs, see Restricting access to an Amazon S3 origin in the *Amazon CloudFront Developer Guide*.

In certain cases, the **Sharing** area also displays values for **Unknown**. If these values appear, Macie wasn't able to determine whether the specified number and percentage of buckets are shared with other accounts, CloudFront OAIs, or CloudFront OACs. For example, a temporary issue or the buckets' permissions settings prevented Macie from retrieving the requisite data. Or Macie wasn't able to fully evaluate the buckets' policies or ACLs. This can also be the case for buckets that exceed the quota for preventative control monitoring. Macie evaluates and monitors the security and privacy of no more than 10,000 buckets for an account—the 10,000 buckets that were most recently created or changed.

Analyzing your Amazon S3 security posture with Macie

To help you perform in-depth analysis and evaluate the security posture of your Amazon Simple Storage Service (Amazon S3) data, Amazon Macie generates and maintains an inventory of your S3 general purpose buckets in each AWS Region where you use Macie. To learn how Macie maintains this inventory for you, see How Macie monitors Amazon S3 data security. If you're the Macie administrator for an organization, the inventory includes data for S3 buckets that your member accounts own.

By using this inventory, you can review your Amazon S3 data estate, and examine details and statistics for key security settings and metrics that apply to individual S3 buckets. For example, you can access breakdowns of each bucket's public access and encryption settings, and the size and number of objects that Macie can analyze to detect sensitive data in each bucket. You can also determine whether you configured sensitive data discovery jobs or automated sensitive data discovery to analyze objects in a bucket. If you have, your inventory data indicates when that analysis most recently occurred. If automated sensitive data discovery is enabled, you can also use the inventory to review the results of automated sensitive data discovery activities that Macie has performed thus far for your Amazon S3 data. For more information, see Discovering sensitive data.

You can browse and filter inventory data by using the **S3 buckets** page on the Amazon Macie console. You can also access your inventory data programmatically by using the DescribeBuckets operation of the Amazon Macie API.

Topics

Reviewing your S3 bucket inventory in Macie

On the Amazon Macie console, the **S3 buckets** page provides detailed insight into the security and privacy of your Amazon Simple Storage Service (Amazon S3) data in the current AWS Region. With this page, you can review and analyze an inventory of your S3 general purpose buckets in the Region, and review detailed information and statistics for individual buckets. For information about how Macie generates and maintains this inventory, see How Macie monitors Amazon S3 data security. If you're the Macie administrator for an organization, your inventory includes details and statistics for S3 buckets that your member accounts own.

The **S3 buckets** page also indicates when Macie most recently retrieved bucket or object metadata from Amazon S3 for your account. You can find this information in the **Last updated** field at the top of the page. If you're the Macie administrator for an organization, this field indicates the earliest date and time when Macie retrieved the data for an account in your organization. For more information, see Data refreshes.

Note that inventory data and statistics don't include data about S3 directory buckets, only general purpose buckets. Macie doesn't monitor or analyze directory buckets. In addition, Macie maintains complete inventory data for no more than 10,000 general purpose buckets for an account. If your

account exceeds this quota, Macie provides complete inventory data for the 10,000 buckets that were most recently created or changed. For all other buckets, Macie provides only a subset of information about each bucket. If you're the Macie administrator for an organization, this quota applies to each account in your organization, not your organization overall.

Also note that most inventory data is limited to buckets that Macie is allowed to access for your account. If a bucket's permissions settings prevent Macie from retrieving information about the bucket or the bucket's objects, Macie can only provide a subset of information about the bucket. If this is the case for a particular bucket, Macie displays a warning icon (⚠) and message for the bucket in your bucket inventory. For the bucket's details, Macie provides data for only a subset of fields: the account ID for the AWS account that owns the bucket; the bucket's name, Amazon Resource Name (ARN), creation date, and Region; and, when Macie most recently retrieved both bucket and object metadata for the bucket as part of the daily refresh cycle. To investigate the issue, review the bucket's policy and permissions settings in Amazon S3. For example, the bucket might have a restrictive bucket policy. For more information, see Allowing Macie to access S3 buckets and objects.

If you prefer to access and query your inventory data programmatically, you can use the DescribeBuckets operation of the Amazon Macie API.

Topics

Reviewing your S3 bucket inventory

The **S3 buckets** page on the Amazon Macie console provides information about your S3 general purpose buckets in the current AWS Region. On this page, a table displays summary information for each bucket in your inventory. To customize your view, you can sort and filter the table. If you choose a bucket in the table, the details panel displays additional information about the bucket. This includes details and statistics for settings and metrics that provide insight into the security and privacy of the bucket's data. You can optionally export data from the table to a comma-separated values (CSV) file.

If automated sensitive data discovery is enabled, you also have the option of reviewing your inventory by using an interactive heat map. The map provides a visual representation of data

sensitivity across your Amazon S3 data estate. It captures the results of automated sensitive data discovery activities that Macie has performed thus far. To learn about this map, see Visualizing data sensitivity with the S3 buckets map.

To review your S3 bucket inventory

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **S3 buckets**. The **S3 buckets** page displays your bucket inventory. If the page displays an interactive map of your inventory, choose table () at the top of the page. Macie then displays the number of buckets in your inventory and a table of the buckets.

   If automated sensitive data discovery is enabled, the default view doesn't display data for buckets that are currently excluded from automated discovery. To display this data, choose **X** in the **Is monitored by automated discovery** filter token below the filter box.
3. At the top of the page, optionally choose refresh () to retrieve the latest bucket metadata from Amazon S3.

   If the information icon () appears next to any bucket names, we recommend that you do this. This icon indicates that a bucket was created during the past 24 hours, possibly after Macie last retrieved bucket and object metadata from Amazon S3 as part of the daily refresh cycle.
4. In the **S3 buckets** table, review a subset of information about each bucket in your inventory:

   - **Sensitivity** – The bucket's current sensitivity score, if automated sensitive data discovery is enabled. For information about the range of sensitivity scores that Macie defines, see Sensitivity scoring for S3 buckets.
   - **Bucket** – The name of the bucket.
   - **Account** – The account ID for the AWS account that owns the bucket.
   - **Classifiable objects** – The total number of objects that Macie can analyze to detect sensitive data in the bucket.
   - **Classifiable size** – The total storage size of all the objects that Macie can analyze to detect sensitive data in the bucket.

Note that this value doesn't reflect the actual size of any compressed objects after they're decompressed. Also, if versioning is enabled for the bucket, this value is based on the storage size of the latest version of each object in the bucket.

- **Monitored by job** – Whether you configured any sensitive data discovery jobs to periodically analyze objects in the bucket on a daily, weekly, or monthly basis.

  If the value for this field is *Yes*, the bucket is explicitly included in a periodic job or the bucket matched the criteria for a periodic job within the past 24 hours. In addition, the status of at least one of those jobs is not *Cancelled*. Macie updates this data on a daily basis.

- **Latest job run** – If you configured any periodic or one-time sensitive data discovery jobs to analyze objects in the bucket, this field indicates the most recent date and time when one of those jobs started to run. Otherwise, a dash (–) appears in this field.

In the preceding data, objects are *classifiable* if they use a supported Amazon S3 storage class and they have a file name extension for a supported file or storage format. You can detect sensitive data in the objects by using Macie. For more information, see Supported storage classes and formats.

5. To analyze your inventory by using the table, do any of the following:

   - To sort the table by a specific field, choose the column heading for the field. To change the sort order, choose the column heading again.
   - To filter the table and display only those buckets that have a specific value for a field, place your cursor in the filter box, and then add a filter condition for the field. To further refine the results, add filter conditions for additional fields. For more information, see Filtering your S3 bucket inventory.

6. To review details and statistics for a particular bucket, choose the bucket's name in the table, and then refer to the details panel.

> **Tip**
>
> You can pivot and drill down on many of the fields in the bucket details panel. To show buckets that have the same value for a field, choose
> 
> in the field. To show buckets that have other values for a field, choose

in the field.

7. To export data from the table to a CSV file, select the checkbox for each row that you want to export, or select the checkbox in the selection column heading to select all rows. Then choose **Export to CSV** at the top of the page. You can export up to 50,000 rows from the table.

Reviewing the details of S3 buckets

To review details and statistics for an S3 general purpose bucket, you can use the details panel on the **S3 buckets** page of the Amazon Macie console. The panel displays details and statistics that provide insight into the security and privacy of a bucket's data.

For example, you can review breakdowns of an S3 bucket's public access settings, and determine whether a bucket is configured to replicate objects or is shared with other AWS accounts. You can also determine whether you configured any sensitive data discovery jobs to inspect the bucket for sensitive data. If you have, you can access details about the job that ran most recently, and optionally display any findings that the job produced.

If automated sensitive data discovery is enabled, you can also use the details panel to review sensitive data discovery statistics and other information about individual S3 buckets. The panel captures the results of automated sensitive data discovery activities that Macie has performed thus far for a bucket. To learn about these details, see Reviewing data sensitivity details for S3 buckets.

To review the details of an S3 bucket

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **S3 buckets**. The **S3 buckets** page displays your bucket inventory.

   If automated sensitive data discovery is enabled, the default view doesn't display data for buckets that are currently excluded from automated discovery. To display this data, choose **X** in the **Is monitored by automated discovery** filter token below the filter box.
3. At the top of the page, optionally choose refresh ( ) to retrieve the latest bucket metadata from Amazon S3.

4. Choose the bucket whose details you want to review. The details panel displays statistics and other information about the bucket.

In the details panel, statistics and information are organized into the following primary sections:

Overview | **Object statistics** | **Server-side encryption** | **Sensitive data discovery** | **Public access** | **Replication** | **Tags**

As you review the information in each section, you can optionally pivot and drill down on certain fields. To show buckets that have the same value for a field, choose  in the field. To show buckets that have other values for a field, choose  in the field.

Overview

This section provides general information about the bucket, such as the bucket's name, when the bucket was created, and the account ID for the AWS account that owns the bucket. Of special note, the **Last updated** field indicates when Macie most recently retrieved metadata from Amazon S3 for the bucket or the bucket's objects.

The **Shared access** field indicates whether the bucket is shared with another AWS account, an Amazon CloudFront origin access identity (OAI), or a CloudFront origin access control (OAC):

- **External** – The bucket is shared with one or more of the following or any combination of the following: a CloudFront OAI, a CloudFront OAC, or an account that's external to (not part of) your organization.
- **Internal** – The bucket is shared with one or more accounts that are internal to (part of) your organization. It isn't shared with a CloudFront OAI or OAC.
- **Not shared** – The bucket isn't shared with another account, a CloudFront OAI, or a CloudFront OAC.
- **Unknown** – Macie wasn't able to evaluate the shared access settings for the bucket. For example, a quota or temporary issue prevented Macie from retrieving and evaluating the requisite data.

To determine whether a bucket is shared with another AWS account, Macie analyzes the bucket policy and access control list (ACL) for the bucket. The analysis is limited to bucket-level settings.

It doesn't reflect any object-level settings for sharing specific objects in the bucket. In addition, an *organization* is defined as a set of Macie accounts that are centrally managed as a group of related accounts through AWS Organizations or by Macie invitation. To learn about Amazon S3 options for sharing buckets, see Access control in the *Amazon Simple Storage Service User Guide*.

> **Note**
>
> In certain cases, Macie might incorrectly indicate that a bucket is shared with an AWS account that's external to (not part of) your organization. This can occur if Macie isn't able to fully evaluate the relationship between the `Principal` element in the bucket's policy and certain AWS global condition context keys or Amazon S3 condition keys in the `Condition` element of the policy. This can be the case for the following condition keys: `aws:PrincipalAccount`, `aws:PrincipalArn`, `aws:PrincipalOrgID`, `aws:PrincipalOrgPaths`, `aws:PrincipalTag`, `aws:PrincipalType`, `aws:SourceAccount`, `aws:SourceArn`, `aws:SourceIp`, `aws:SourceOrgID`, `aws:SourceOrgPaths`, `aws:SourceVpc`, `aws:SourceVpce`, `aws:userid`, `s3:DataAccessPointAccount`, and `s3:DataAccessPointArn`. We recommend that you review the bucket's policy to determine whether this access is intended and safe.

To determine whether a bucket is shared with a CloudFront OAI or OAC, Macie analyzes the bucket policy for the bucket. A CloudFront OAI or OAC allows users to access a bucket's objects through one or more specified CloudFront distributions. To learn about CloudFront OAIs and OACs, see Restricting access to an Amazon S3 origin in the *Amazon CloudFront Developer Guide*.

The **Overview** section also includes the **Latest automated discovery run** field. This field indicates when Macie most recently analyzed objects in the bucket while performing automated sensitive data discovery. If this analysis hasn't occurred, a dash (–) appears in this field.

Object statistics

This section provides information about the objects in the bucket, starting with the total number of objects in the bucket (**Total count**), the total storage size of all those objects (**Total storage size**), and the total storage size of all the objects that are compressed (.gz, .gzip, or .zip) files (**Total compressed size**). Additional statistics in this section can help you assess how much data Macie can analyze to detect sensitive data in the bucket.

If you recently created the bucket or made significant changes to the bucket's objects during the past 24 hours, optionally choose refresh () to retrieve the latest metadata for the bucket's objects. Macie displays the information icon () to help you determine whether this might be the case. The refresh option is available if a bucket stores 30,000 or fewer objects.

As you review the statistics in this section, keep the following in mind:

- If versioning is enabled for the bucket, size values are based on the storage size of the latest version of each object in the bucket.
- If the bucket stores compressed objects, size values don't reflect the actual size of those objects after they're decompressed.
- If you refresh object metadata for a bucket, Macie temporarily reports *Unknown* for encryption statistics that apply to the objects. Macie will re-evaluate and update the data for these statistics when it performs the next daily refresh of bucket and object metadata, which is within 24 hours.
- By default, object counts and size values include data for any object parts that the bucket contains as a result of incomplete multipart uploads. If you refresh object metadata for a bucket, Macie excludes data for object parts from the recalculated values. When Macie performs the next daily refresh of bucket and object metadata (within 24 hours), Macie recalculates and updates the values for these statistics and includes data for object parts in the values again.

  Note that Macie can't analyze object parts to detect sensitive data. Amazon S3 must first finish assembling the parts into one or more objects for Macie to analyze. For information about multipart uploads and object parts, including how to delete parts automatically with lifecycle rules, see Uploading and copying objects using multipart upload in the *Amazon Simple Storage Service User Guide*. To identify buckets that contain object parts, you can refer to *incomplete multipart upload* metrics in Amazon S3 Storage Lens. For more information, see Assessing your storage activity and usage in the *Amazon Simple Storage Service User Guide*.

Object statistics are organized as follows.

Classifiable objects

This section indicates the total number of objects that Macie can analyze to detect sensitive data and the total storage size of those objects. These objects use a supported Amazon S3

storage class and have a file name extension for a supported file or storage format. You can detect sensitive data in the objects by using Macie. For more information, see Supported storage classes and formats.

Unclassifiable objects

This section indicates the total number of objects that Macie can't analyze to detect sensitive data and the total storage size of those objects. These objects don't use a supported Amazon S3 storage class or they don't have a file name extension for a supported file or storage format.

Unclassifiable objects: Storage class

This section provides a breakdown of the number and storage size of the objects that Macie can't analyze because the objects don't use a supported Amazon S3 storage class.

Unclassifiable objects: File type

This section provides a breakdown of the number and storage size of the objects that Macie can't analyze because the objects don't have a file name extension for a supported file or storage format.

Objects by encryption type

This section provides a breakdown of the number of objects that use each type of encryption that Amazon S3 supports:

- **Customer provided** – The number of objects that are encrypted with a customer-provided key. These objects use SSE-C encryption.
- **AWS KMS managed** – The number of objects that are encrypted with an AWS KMS key, either an AWS managed key or a customer managed key. These objects use DSSE-KMS or SSE-KMS encryption.
- **Amazon S3 managed** – The number of objects that are encrypted with an Amazon S3 managed key. These objects use SSE-S3 encryption.
- **No encryption** – The number of objects that aren't encrypted or use client-side encryption. (If an object is encrypted using client-side encryption, Macie can't access and report encryption data for the object.)
- **Unknown** – The number of objects that Macie doesn't have current encryption metadata for. This typically occurs if you recently chose to manually refresh the metadata for the bucket's objects. Macie will update the encryption statistics when it performs the next daily refresh of bucket and object metadata, which is within 24 hours.

For information about each supported encryption type, see Protecting data with encryption in the *Amazon Simple Storage Service User Guide*.

Server-side encryption

This section provides insight into the server-side encryption settings for the bucket.

The **Encryption required by bucket policy** field indicates whether the bucket's policy requires server-side encryption of objects when objects are added to the bucket:

- **No** – The bucket doesn't have a bucket policy or the bucket's policy doesn't require server-side encryption of new objects. If a bucket policy exists, it doesn't require PutObject requests to include a valid server-side encryption header.
- **Yes** – The bucket's policy requires server-side encryption of new objects. **PutObject** requests for the bucket must include a valid server-side encryption header. Otherwise, Amazon S3 denies the request.
- **Unknown** – Macie wasn't able to evaluate the bucket's policy to determine whether it requires server-side encryption of new objects. For example, a quota or issue prevented Macie from retrieving and evaluating the policy.

For this assessment, valid server-side encryption headers are: `x-amz-server-side-encryption` with a value of `AES256` or `aws:kms`, and `x-amz-server-side-encryption-customer-algorithm` with a value of `AES256`. For information about using bucket policies to require server-side encryption of new objects, see Protecting data with server-side encryption in the *Amazon Simple Storage Service User Guide*.

The **Default encryption** field indicates which server-side encryption algorithm the bucket is configured to apply by default to objects that are added to the bucket:

- **AES256** – The bucket's default encryption settings are configured to encrypt new objects with an Amazon S3 managed key. New objects are encrypted automatically using SSE-S3 encryption.
- **aws:kms** – The bucket's default encryption settings are configured to encrypt new objects with an AWS KMS key, either an AWS managed key or a customer managed key. New objects are encrypted automatically using SSE-KMS encryption. The **AWS KMS key** field shows the Amazon Resource Name (ARN) or unique identifier (key ID) for the key that's used.
- **aws:kms:dsse** – The bucket's default encryption settings are configured to encrypt new objects with an AWS KMS key, either an AWS managed key or a customer managed key. New objects are

encrypted automatically using DSSE-KMS encryption. The **AWS KMS key** field shows the ARN or key ID for the key that's used.

- **None** – The bucket's default encryption settings don't specify server-side encryption behavior for new objects.

Starting January 5, 2023, Amazon S3 automatically applies server-side encryption with Amazon S3 managed keys (SSE-S3) as the base level of encryption for objects that are added to buckets. You can optionally configure a bucket's default encryption settings to instead use server-side encryption with an AWS KMS key (SSE-KMS) or dual-layer server-side encryption with an AWS KMS key (DSSE-KMS). For information about default encryption settings and options, see Setting default server-side encryption behavior for S3 buckets in the *Amazon Simple Storage Service User Guide*.

Sensitive data discovery

This section indicates whether you configured any sensitive data discovery jobs to periodically analyze objects in the bucket on a daily, weekly, or monthly basis. If the value for the **Actively monitored by job** field is *Yes*, the bucket is explicitly included in a periodic job or the bucket matched the criteria for a periodic job within the past 24 hours. In addition, the status of at least one of those jobs is not *Cancelled*. Macie updates this data on a daily basis.

If you configured any type of sensitive data discovery job (either a periodic job or a one-time job) to analyze objects in the bucket, the **Latest job** field provides the unique identifier for the job that most recently started to run. The **Latest job run** field indicates when that job started to run.

> **Tip**
>
> To display all the sensitive data findings that the job produced, choose the link in the **Latest job** field. In the job details panel that appears, choose **Show results** at the top of the panel, and then choose **Show findings**.

Public access

This section indicates whether the bucket is publicly accessible. It also provides a breakdown of the various account- and bucket-level settings that determine whether this is the case. The **Effective permission** field indicates the cumulative result of these settings:

- **Not public** – The bucket isn't publicly accessible.

- **Public** – The bucket is publicly accessible.
- **Unknown** – Macie wasn't able to evaluate all the public access settings for the bucket. For example, a quota or temporary issue prevented Macie from retrieving and evaluating the requisite data.

For this evaluation, Macie analyzes a combination of account- and bucket-level settings for each bucket: the block public access settings for the account; the block public access settings for the bucket; the bucket policy for the bucket; and, the access control list (ACL) for the bucket. Note that the evaluation doesn't include object-level settings that enable public access to specific objects in a bucket.

To learn about Amazon S3 settings for managing public access to buckets and bucket data, see Access control and Blocking public access to your Amazon S3 storage in the *Amazon Simple Storage Service User Guide*.

Replication

In this section, the **Replicated** field indicates whether the bucket is configured to replicate objects to other buckets. If the value for this field is *Yes*, one or more replication rules are configured and enabled for the bucket. This section then also lists the account ID for each AWS account that owns a destination bucket.

The **Replicated externally** field indicates whether the bucket is configured to replicate objects to buckets for AWS accounts that are external to (not part of) your organization. An *organization* is a set of Macie accounts that are centrally managed as a group of related accounts through AWS Organizations or by Macie invitation. If the value for this field is *Yes*, a replication rule is configured and enabled for the bucket, and the rule is configured to replicate objects to a bucket that's owned by an external AWS account.

> **Note**
>
> Under certain conditions, Macie might incorrectly indicate that a bucket is configured to replicate objects to a bucket that's owned by an external AWS account. This can occur if the destination bucket was created in a different AWS Region during the preceding 24 hours, after Macie retrieved bucket and object metadata from Amazon S3 as part of the daily refresh cycle. To investigate the issue by using Macie, choose refresh () to retrieve the latest bucket metadata from Amazon S3. Then review the list of account IDs

in this section. For deeper investigation, use Amazon S3 to review the replication rules for the bucket.

To learn about Amazon S3 options and settings for replicating bucket objects, see Replicating objects in the *Amazon Simple Storage Service User Guide*.

Tags

If tags are associated with the bucket, this section appears in the panel and lists those tags. Tags are labels that you can define and assign to certain types of AWS resources, including S3 buckets. Each tag consists of a required tag key and an optional tag value.

To learn about tagging buckets, see Using cost allocation S3 bucket tags in the *Amazon Simple Storage Service User Guide*.

Filtering your S3 bucket inventory in Macie

To identify and focus on buckets that have specific characteristics, you can filter your S3 bucket inventory on the Amazon Macie console and in queries that you submit programmatically using the Amazon Macie API. When you create a filter, you use specific bucket attributes to define criteria for including or excluding buckets from a view or from query results. A *bucket attribute* is a field that stores specific metadata for a bucket.

In Macie, a filter consists of one or more conditions. Each condition, also referred to as a *criterion*, consists of three parts:

- An attribute-based field, such as **Bucket name**, **Tag key**, or **Defined in job**.
- An operator, such as *equals* or *not equals*.
- One or more values. The type and number of values depends on the field and operator that you choose.

How you define and apply filter conditions depends on whether you use the Amazon Macie console or the Amazon Macie API.

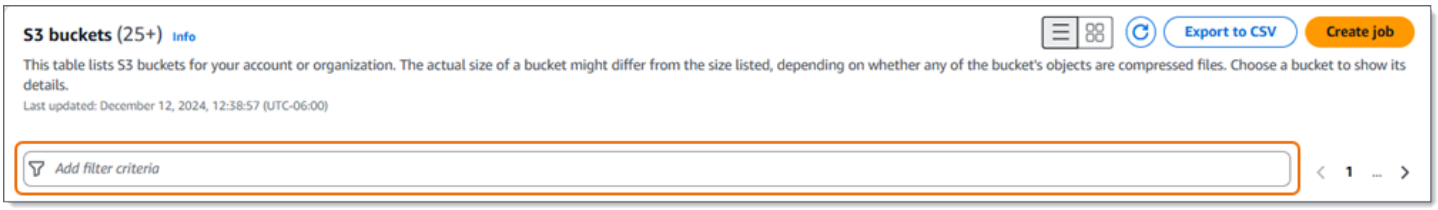
Topics

- Filtering your inventory on the Amazon Macie console

- Filtering your inventory programmatically with the Amazon Macie API

Filtering your inventory on the Amazon Macie console

If you use the Amazon Macie console to filter your S3 bucket inventory, Macie provides options to help you choose fields, operators, and values for individual conditions. You access these options by using the filter box on the **S3 buckets** page, as shown in the following image.

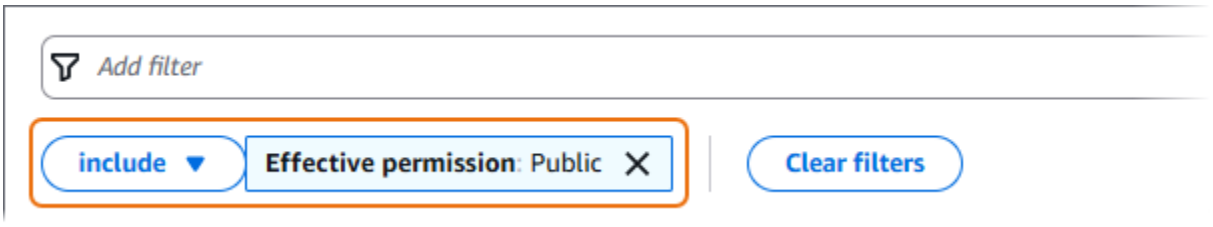


When you place your cursor in the filter box, Macie displays a list of fields that you can use in filter conditions. The fields are organized by logical category. For example, the **Common fields** category includes fields that store general information about an S3 bucket. **Public access** categories include fields that store data about the various types of public access settings that can apply to a bucket. The fields are sorted alphabetically within each category.

To add a condition, start by choosing a field from the list. To find a field, browse the complete list, or enter part of the field's name to narrow the list of fields.

Depending on the field that you choose, Macie displays different options. The options reflect the type and nature of the field that you choose. For example, if you choose the **Shared access** field, Macie displays a list of values to choose from. If you choose the **Bucket name** field, Macie displays a text box in which you can enter the name of an S3 bucket. Whichever field you choose, Macie guides you through the steps to add a condition that includes the required settings for the field.

After you add a condition, Macie applies the criteria for the condition and displays the condition in a filter token below the filter box, as shown in the following image.



In this example, the condition is configured to include all buckets that are publicly accessible, and to exclude all other buckets. It returns buckets where the value for the **Effective permission** field *equals* **Public**.

As you add more conditions, Macie applies their criteria and displays them below the filter box. If you add multiple conditions, Macie uses AND logic to join the conditions and evaluate the filter criteria. This means that an S3 bucket matches the filter criteria only if it matches all the conditions in the filter. You can refer to the area below the filter box at any time to determine which criteria you've applied.

To filter your inventory by using the console

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **S3 buckets**. The **S3 buckets** page displays your bucket inventory.

   If automated sensitive data discovery is enabled, the default view doesn't display data for buckets that are currently excluded from automated discovery. If you're the Macie administrator for an organization, it also doesn't display data for accounts that automated discovery is currently disabled for. To display this data, choose **X** in the **Is monitored by automated discovery** filter token below the filter box.
3. At the top of the page, optionally choose refresh () to retrieve the latest bucket metadata from Amazon S3.
4. Place your cursor in the filter box, and then choose the field to use for the condition.
5. Choose or enter the appropriate type of value for the field, keeping the following tips in mind.

   Dates, times, and time ranges

   For dates and times, use the **From** and **To** boxes to define an inclusive time range:

   - To define a fixed time range, use the **From** and **To** boxes to specify the first date and time and the last date and time in the range, respectively.
   - To define a relative time range that starts at a certain date and time and ends at the current time, enter the start date and time in the **From** boxes, and delete any text in **To** boxes.
   - To define a relative time range that ends at a certain date and time, enter the end date and time in the **To** boxes, and delete any text in the **From** boxes.

   Note that time values use 24-hour notation. If you use the date picker to choose dates, you can refine the values by entering text directly in the **From** and **To** boxes.

Numbers and numeric ranges

For numeric values, use the **From** and **To** boxes to enter integers that define an inclusive numeric range:

- To define a fixed numeric range, use the **From** and **To** boxes to specify the lowest and highest numbers in the range, respectively.
- To define a fixed numeric range that's limited to one specific value, enter the value in both the **From** and **To** boxes. For example, to include only those S3 buckets that store exactly 15 objects, enter **15** in the **From** and **To** boxes.
- To define a relative numeric range that starts at a certain number, enter the number in the **From** box, and don't enter any text in the **To** box.
- To define a relative numeric range that ends at a certain number, enter the number in the **To** box, and don't enter any text in the **From** box.

Text (string) values

For this type of value, enter a complete, valid value for the field. Values are case sensitive.

Note that you can't use a partial value or wildcard characters in this type of value. The only exception is the **Bucket name** field. For that field, you can specify a prefix instead of a complete bucket name. For example, to find all S3 buckets whose names begin with *my-S3*, enter **`my-S3`** as the filter value for **Bucket name** field. If you enter any other value, such as **`My-s3`** or **`my*`**, Macie won't return the buckets.

6. When you finish adding a value for the field, choose **Apply**. Macie applies the filter criteria and displays the condition in a filter token below the filter box.
7. Repeat steps 4 through 6 for each additional condition that you want to add.
8. To remove a condition, choose the **X** in the filter token for the condition.
9. To change a condition, remove the condition by choosing the **X** in the filter token for the condition. Then repeat steps 4 through 6 to add a condition with the correct settings.

Filtering your inventory programmatically with the Amazon Macie API

To filter your S3 bucket inventory programmatically, specify filter criteria in queries that you submit using the DescribeBuckets operation of the Amazon Macie API. This operation returns an array of objects. Each object contains statistical data and other information about a bucket that matches the filter criteria.

To specify filter criteria in a query, include a map of filter conditions in your request. For each condition, specify a field, an operator, and one or more values for the field. The type and number of values depends on the field and operator that you choose. For information about the fields, operators, and types of values that you can use in a condition, see Amazon S3 Data Sources in the *Amazon Macie API Reference*.

The following examples show you how to specify filter criteria in queries that you submit using the AWS Command Line Interface (AWS CLI). You can also do this by using a current version of another AWS command line tool or an AWS SDK, or by sending HTTPS requests directly to Macie. For information about AWS tools and SDKs, see Tools to Build on AWS.

Examples

- Example: Find buckets by bucket name
- Example: Find buckets that are publicly accessible
- Example: Find buckets that store unencrypted objects
- Example: Find buckets that replicate data to external accounts
- Example: Find buckets that aren't monitored by a sensitive data discovery job
- Example: Find buckets that aren't monitored by automated sensitive data discovery
- Example: Find buckets based on multiple criteria

The examples use the describe-buckets command. If the command runs successfully, Macie returns a `buckets` array. The array contains an object for each bucket that's in the current AWS Region and matches the filter criteria. For an example of this output, expand the following section.

Example of a `buckets` array

In this example, the `buckets` array provides details about two buckets that match the filter criteria specified in a query.

```
{
    "buckets": [
        {
            "accountId": "123456789012",
            "allowsUnencryptedObjectUploads": "FALSE",
            "automatedDiscoveryMonitoringStatus": "MONITORED",
            "bucketArn": "arn:aws:s3:::amzn-s3-demo-bucket1",
            "bucketCreatedAt": "2020-05-18T19:54:00+00:00",
            "bucketName": "amzn-s3-demo-bucket1",
```

```
            "classifiableObjectCount": 13,
            "classifiableSizeInBytes": 1592088,
            "jobDetails": {
                "isDefinedInJob": "TRUE",
                "isMonitoredByJob": "TRUE",
                "lastJobId": "08c81dc4a2f3377fae45c9ddaexample",
                "lastJobRunTime": "2024-05-26T14:55:30.270000+00:00"
            },
            "lastAutomatedDiscoveryTime": "2024-06-07T19:11:25.364000+00:00",
            "lastUpdated": "2024-06-12T07:33:06.337000+00:00",
            "objectCount": 13,
            "objectCountByEncryptionType": {
                "customerManaged": 0,
                "kmsManaged": 2,
                "s3Managed": 7,
                "unencrypted": 4,
                "unknown": 0
            },
            "publicAccess": {
                "effectivePermission": "NOT_PUBLIC",
                "permissionConfiguration": {
                    "accountLevelPermissions": {
                        "blockPublicAccess": {
                            "blockPublicAcls": true,
                            "blockPublicPolicy": true,
                            "ignorePublicAcls": true,
                            "restrictPublicBuckets": true
                        }
                    },
                    "bucketLevelPermissions": {
                        "accessControlList": {
                            "allowsPublicReadAccess": false,
                            "allowsPublicWriteAccess": false
                        },
                        "blockPublicAccess": {
                            "blockPublicAcls": true,
                            "blockPublicPolicy": true,
                            "ignorePublicAcls": true,
                            "restrictPublicBuckets": true
                        },
                        "bucketPolicy": {
                            "allowsPublicReadAccess": false,
                            "allowsPublicWriteAccess": false
                        }
```

```
                    }
                }
            },
            "region": "us-east-1",
            "replicationDetails": {
                "replicated": false,
                "replicatedExternally": false,
                "replicationAccounts": []
            },
            "sensitivityScore": 78,
            "serverSideEncryption": {
                "kmsMasterKeyId": null,
                "type": "NONE"
            },
            "sharedAccess": "NOT_SHARED",
            "sizeInBytes": 4549746,
            "sizeInBytesCompressed": 0,
            "tags": [
                {
                    "key": "Division",
                    "value": "HR"
                },
                {
                    "key": "Team",
                    "value": "Recruiting"
                }
            ],
            "unclassifiableObjectCount": {
                "fileType": 0,
                "storageClass": 0,
                "total": 0
            },
            "unclassifiableObjectSizeInBytes": {
                "fileType": 0,
                "storageClass": 0,
                "total": 0
            },
            "versioning": true
        },
        {
            "accountId": "123456789012",
            "allowsUnencryptedObjectUploads": "TRUE",
            "automatedDiscoveryMonitoringStatus": "MONITORED",
            "bucketArn": "arn:aws:s3:::amzn-s3-demo-bucket2",
```

```
            "bucketCreatedAt": "2020-11-25T18:24:38+00:00",
            "bucketName": "amzn-s3-demo-bucket2",
            "classifiableObjectCount": 8,
            "classifiableSizeInBytes": 133810,
            "jobDetails": {
                "isDefinedInJob": "TRUE",
                "isMonitoredByJob": "FALSE",
                "lastJobId": "188d4f6044d621771ef7d65f2example",
                "lastJobRunTime": "2024-04-09T19:37:11.511000+00:00"
            },
            "lastAutomatedDiscoveryTime": "2024-06-07T19:11:25.364000+00:00",
            "lastUpdated": "2024-06-12T07:33:06.337000+00:00",
            "objectCount": 8,
            "objectCountByEncryptionType": {
                "customerManaged": 0,
                "kmsManaged": 0,
                "s3Managed": 8,
                "unencrypted": 0,
                "unknown": 0
            },
            "publicAccess": {
                "effectivePermission": "NOT_PUBLIC",
                "permissionConfiguration": {
                    "accountLevelPermissions": {
                        "blockPublicAccess": {
                            "blockPublicAcls": true,
                            "blockPublicPolicy": true,
                            "ignorePublicAcls": true,
                            "restrictPublicBuckets": true
                        }
                    },
                    "bucketLevelPermissions": {
                        "accessControlList": {
                            "allowsPublicReadAccess": false,
                            "allowsPublicWriteAccess": false
                        },
                        "blockPublicAccess": {
                            "blockPublicAcls": true,
                            "blockPublicPolicy": true,
                            "ignorePublicAcls": true,
                            "restrictPublicBuckets": true
                        },
                        "bucketPolicy": {
                            "allowsPublicReadAccess": false,
```

```
                            "allowsPublicWriteAccess": false
                        }
                    }
                }
            },
            "region": "us-east-1",
            "replicationDetails": {
                "replicated": false,
                "replicatedExternally": false,
                "replicationAccounts": []
            },
            "sensitivityScore": 95,
            "serverSideEncryption": {
                "kmsMasterKeyId": null,
                "type": "AES256"
            },
            "sharedAccess": "EXTERNAL",
            "sizeInBytes": 175978,
            "sizeInBytesCompressed": 0,
            "tags": [
                {
                    "key": "Division",
                    "value": "HR"
                },
                {
                    "key": "Team",
                    "value": "Recruiting"
                }
            ],
            "unclassifiableObjectCount": {
                "fileType": 3,
                "storageClass": 0,
                "total": 3
            },
            "unclassifiableObjectSizeInBytes": {
                "fileType": 2999826,
                "storageClass": 0,
                "total": 2999826
            },
            "versioning": true
        }
    ]
}
```

If no buckets match the filter criteria, Macie returns an empty `buckets` array.

```
{
    "buckets": []
}
```

Example: Find buckets by bucket name

This example queries metadata for buckets that are in the current AWS Region and have names beginning with *my-S3*.

For Linux, macOS, or Unix:

```
$ aws macie2 describe-buckets --criteria '{"bucketName":{"prefix":"my-S3"}}'
```

For Microsoft Windows:

```
C:\> aws macie2 describe-buckets --criteria={\"bucketName\":{\"prefix\":\"my-S3\"}}
```

Where:

- *`bucketName`* specifies the JSON name of the **Bucket name** field.
- *`prefix`* specifies the *prefix* operator.
- *`my-S3`* is the value for the **Bucket name** field.

Example: Find buckets that are publicly accessible

This example queries metadata for buckets that are in the current AWS Region and, based on a combination of permissions settings, are publicly accessible.

For Linux, macOS, or Unix:

```
$ aws macie2 describe-buckets --criteria '{"publicAccess.effectivePermission":{"eq":
["PUBLIC"]}}'
```

For Microsoft Windows:

```
C:\> aws macie2 describe-buckets --criteria={\"publicAccess.effectivePermission\":
{\"eq\":[\"PUBLIC\"]}}
```

Where:

- *publicAccess.effectivePermission* specifies the JSON name of the **Effective permission** field.
- *eq* specifies the *equals* operator.
- *PUBLIC* is an enumerated value for the **Effective permission** field.

Example: Find buckets that store unencrypted objects

This example queries metadata for buckets that are in the current AWS Region and store unencrypted objects.

For Linux, macOS, or Unix:

```
$ aws macie2 describe-buckets --criteria '{"objectCountByEncryptionType.unencrypted":
{"gte":1}}'
```

For Microsoft Windows:

```
C:\> aws macie2 describe-buckets --
criteria={\"objectCountByEncryptionType.unencrypted\":{\"gte\":1}}
```

Where:

- *objectCountByEncryptionType.unencrypted* specifies the JSON name of the **No encryption** field.
- *gte* specifies the *greater than or equal to* operator.
- *1* is the lowest value in an inclusive, relative numeric range for the **No encryption** field.

Example: Find buckets that replicate data to external accounts

This example queries metadata for buckets that are in the current AWS Region and are configured to replicate objects to buckets for an AWS account that isn't part of your organization.

For Linux, macOS, or Unix:

```
$ aws macie2 describe-buckets --criteria '{"replicationDetails.replicatedExternally":
{"eq":["true"]}}'
```

For Microsoft Windows:

```
C:\> aws macie2 describe-buckets --
criteria={\"replicationDetails.replicatedExternally\":{\"eq\":[\"true\"]}}
```

Where:

- *replicationDetails.replicatedExternally* specifies the JSON name of the **Replicated externally** field.
- *eq* specifies the *equals* operator.
- *true* specifies a Boolean value for the **Replicated externally** field.

Example: Find buckets that aren't monitored by a sensitive data discovery job

This example queries metadata for buckets that are in the current AWS Region and aren't associated with any periodic sensitive data discovery jobs.

For Linux, macOS, or Unix:

```
$ aws macie2 describe-buckets --criteria '{"jobDetails.isMonitoredByJob":{"eq":
["FALSE"]}}'
```

For Microsoft Windows:

```
C:\> aws macie2 describe-buckets --criteria={\"jobDetails.isMonitoredByJob\":{\"eq\":
[\"FALSE\"]}}
```

Where:

- *jobDetails.isMonitoredByJob* specifies the JSON name of the **Actively monitored by job** field.
- *eq* specifies the *equals* operator.
- *FALSE* is an enumerated value for the **Actively monitored by job** field.

Example: Find buckets that aren't monitored by automated sensitive data discovery

This example queries metadata for buckets that are in the current AWS Region and are excluded from automated sensitive data discovery.

For Linux, macOS, or Unix:

```
$ aws macie2 describe-buckets --criteria '{"automatedDiscoveryMonitoringStatus":{"eq":
["NOT_MONITORED"]}}'
```

For Microsoft Windows:

```
C:\> aws macie2 describe-buckets --criteria={\"automatedDiscoveryMonitoringStatus\":
{\"eq\":[\"NOT_MONITORED\"]}}
```

Where:

- *automatedDiscoveryMonitoringStatus* specifies the JSON name of the **Is monitored by automated discovery** field.
- *eq* specifies the *equals* operator.
- *NOT_MONITORED* is an enumerated value for the **Is monitored by automated discovery** field.

Example: Find buckets based on multiple criteria

This example queries metadata for buckets that are in the current AWS Region and match the following criteria: are publicly accessible based on a combination of permission settings; store unencrypted objects; and, aren't associated with any periodic sensitive data discovery jobs.

For Linux, macOS, or Unix, using the backslash (\) line-continuation character to improve readability:

```
$ aws macie2 describe-buckets \
--criteria '{"publicAccess.effectivePermission":{"eq":
["PUBLIC"]},"objectCountByEncryptionType.unencrypted":
{"gte":1},"jobDetails.isMonitoredByJob":{"eq":["FALSE"]}}'
```

For Microsoft Windows, using the caret (^) line-continuation character to improve readability:

```
C:\> aws macie2 describe-buckets ^
--criteria={\"publicAccess.effectivePermission\":{\"eq\":
[\"PUBLIC\"]},\"objectCountByEncryptionType.unencrypted\":{\"gte\":1},
\"jobDetails.isMonitoredByJob\":{\"eq\":[\"FALSE\"]}}
```

Where:

- *publicAccess.effectivePermission* specifies the JSON name of the **Effective permission** field, and:
  - *eq* specifies the *equals* operator.
  - *PUBLIC* is an enumerated value for the **Effective permission** field.
- *objectCountByEncryptionType.unencrypted* specifies the JSON name of the **No encryption** field, and:
  - *gte* specifies the *greater than or equal to* operator.
  - *1* is the lowest value in an inclusive, relative numeric range for the **No encryption** field.
- *jobDetails.isMonitoredByJob* specifies the JSON name of the **Actively monitored by job** field, and:
  - *eq* specifies the *equals* operator.
  - *FALSE* is an enumerated value for the **Actively monitored by job** field.

Allowing Macie to access S3 buckets and objects

When you enable Amazon Macie for your AWS account, Macie creates a service-linked role that grants Macie the permissions that it requires to call Amazon Simple Storage Service (Amazon S3) and other AWS services on your behalf. A service-linked role simplifies the process of setting up an AWS service because you don't have to manually add permissions for the service to complete actions on your behalf. To learn about this type of role, see IAM roles in the *AWS Identity and Access Management User Guide*.

The permissions policy for the Macie service-linked role (`AWSServiceRoleForAmazonMacie`) allows Macie to perform actions that include retrieving information about your S3 buckets and objects, and retrieving objects from your buckets. If you're the Macie administrator for an organization, the policy also allows Macie to perform these actions on your behalf for member accounts in your organization.

Macie uses these permissions to perform tasks such as:

- Generate and maintain an inventory of your S3 general purpose buckets.
- Provide statistical and other data about the buckets and objects in the buckets.
- Monitor and evaluate the buckets for security and access control.

- Analyze objects in the buckets to detect sensitive data.

In most cases, Macie has the permissions that it needs to perform these tasks. However, if an S3 bucket has a restrictive bucket policy, the policy might prevent Macie from performing some or all of these tasks.

A *bucket policy* is a resource-based AWS Identity and Access Management (IAM) policy that specifies which actions a principal (user, account, service, or other entity) can perform on an S3 bucket, and the conditions under which a principal can perform those actions. The actions and conditions can apply to bucket-level operations, such as retrieving information about a bucket, and object-level operations, such as retrieving objects from a bucket.

Bucket policies typically grant or restrict access by using explicit `Allow` or `Deny` statements and conditions. For example, a bucket policy might contain an `Allow` or `Deny` statement that denies access to the bucket unless specific source IP addresses, Amazon Virtual Private Cloud (Amazon VPC) endpoints, or VPCs are used to access the bucket. For information about using bucket policies to grant or restrict access to buckets, see Bucket policies for Amazon S3 and How Amazon S3 authorizes a request in the *Amazon Simple Storage Service User Guide*.

If a bucket policy uses an explicit `Allow` statement, the policy doesn't prevent Macie from retrieving information about the bucket and the bucket's objects, or retrieving objects from the bucket. This is because the `Allow` statements in the permissions policy for the Macie service-linked role grant these permissions.

However, if a bucket policy uses an explicit `Deny` statement with one or more conditions, Macie might not be allowed to retrieve information about the bucket or the bucket's objects, or retrieve the bucket's objects. For example, if a bucket policy explicitly denies access from all sources except a specific IP address, Macie won't be allowed to analyze the bucket's objects when you run a sensitive data discovery job. This is because restrictive bucket policies take precedence over the `Allow` statements in the permissions policy for the Macie service-linked role.

To allow Macie to access an S3 bucket that has a restrictive bucket policy, you can add a condition for the Macie service-linked role (`AWSServiceRoleForAmazonMacie`) to the bucket policy. The condition can exclude the Macie service-linked role from matching the `Deny` restriction in the policy. It can do this by using the `aws:PrincipalArn` global condition context key and the Amazon Resource Name (ARN) of the Macie service-linked role.

The following procedure guides you through this process and provides an example.

To add the Macie service-linked role to a bucket policy

1. Sign in to the AWS Management Console and open the Amazon S3 console at https://console.aws.amazon.com/s3/.
2. In the navigation pane, choose **Buckets**.
3. Choose the S3 bucket that you want to allow Macie to access.
4. On the **Permissions** tab, under **Bucket policy**, choose **Edit**.
5. In the **Bucket policy** editor, identify each `Deny` statement that restricts access and prevents Macie from accessing the bucket or the bucket's objects.
6. In each `Deny` statement, add a condition that uses the `aws:PrincipalArn` global condition context key and specifies the ARN of the Macie service-linked role for your AWS account.

   The value for the condition key should be `arn:aws:iam::123456789012:role/aws-service-role/macie.amazonaws.com/AWSServiceRoleForAmazonMacie`, where *123456789012* is the account ID for your AWS account.

Where you add this to a bucket policy depends on the structure, elements, and conditions that the policy currently contains. To learn about supported structures and elements, see Policies and permissions in Amazon S3 in the *Amazon Simple Storage Service User Guide*.

The following is an example of a bucket policy that uses an explicit `Deny` statement to restrict access to an S3 bucket named `amzn-s3-demo-bucket`. With the current policy, the bucket can be accessed only from the VPC endpoint whose ID is `vpce-1a2b3c4d`. Access from all other VPC endpoints is denied, including access from the AWS Management Console and Macie.

JSON

```
{
   "Version": "2012-10-17",
   "Id": "Policy1415115example",
   "Statement": [
      {
         "Sid": "Access only from specific VPCE",
         "Effect": "Deny",
         "Principal": "*",
         "Action": "s3:*",
         "Resource": [
            "arn:aws:s3:::amzn-s3-demo-bucket",
```

```
            "arn:aws:s3:::amzn-s3-demo-bucket/*"
         ],
         "Condition": {
            "StringNotEquals": {
               "aws:SourceVpce": "vpce-1a2b3c4d"
            }
         }
      }
   ]
}
```

To change this policy and allow Macie to access the S3 bucket and the bucket's objects, we can add a condition that uses the `StringNotLike` condition operator and the `aws:PrincipalArn` global condition context key. This additional condition excludes the Macie service-linked role from matching the `Deny` restriction.

JSON

```
{
   "Version": "2012-10-17",
   "Id":" Policy1415115example ",
   "Statement": [
      {
         "Sid": "Access only from specific VPCE and Macie",
         "Effect": "Deny",
         "Principal": "*",
         "Action": "s3:*",
         "Resource": [
            "arn:aws:s3:::amzn-s3-demo-bucket",
            "arn:aws:s3:::amzn-s3-demo-bucket/*"
         ],
         "Condition": {
            "StringNotEquals": {
               "aws:SourceVpce": "vpce-1a2b3c4d"
            },
            "StringNotLike": {
               "aws:PrincipalArn": "arn:aws:iam::123456789012:role/aws-service-
role/macie.amazonaws.com/AWSServiceRoleForAmazonMacie"
            }
         }
      }
```

```
    ]
}
```

In the preceding example, the `StringNotLike` condition operator uses the `aws:PrincipalArn` condition context key to specify the ARN of the Macie service-linked role, where:

- `123456789012` is the account ID for the AWS account that's permitted to use Macie to retrieve information about the bucket and the bucket's objects, and retrieve objects from the bucket.
- `macie.amazonaws.com` is the identifier for the Macie service principal.
- `AWSServiceRoleForAmazonMacie` is the name of the Macie service-linked role.

We used the `StringNotLike` operator because the policy already uses a `StringNotEquals` operator. A policy can use the `StringNotEquals` operator only once.

For additional policy examples and detailed information about managing access to Amazon S3 resources, see Access control in the *Amazon Simple Storage Service User Guide*.

Discovering sensitive data with Macie

With Amazon Macie, you can automate discovery, logging, and reporting of sensitive data in your Amazon Simple Storage Service (Amazon S3) data estate. You can do this in two ways: by configuring Macie to perform automated sensitive data discovery, and by creating and running sensitive data discovery jobs.

Automated sensitive data discovery provides broad visibility into where sensitive data might reside in your Amazon S3 data estate. With this option, Macie evaluates your S3 bucket inventory on a daily basis and uses sampling techniques to identify and select representative S3 objects from your buckets. Macie then retrieves and analyzes the selected objects, inspecting them for sensitive data. For more information, see Performing automated sensitive data discovery.

Sensitive data discovery jobs provide deeper, more targeted analysis. With this option, you define the breadth and depth of the analysis—specific S3 buckets that you select or buckets that match specific criteria. You can also refine the scope of the analysis by choosing options such as custom criteria that derive from properties of S3 objects. In addition, you can configure a job to run only once for on-demand analysis and assessment, or on a recurring basis for periodic analysis, assessment, and monitoring. For more information, see Running sensitive data discovery jobs.

With either option, automated sensitive data discovery or sensitive data discovery jobs, you can configure Macie to analyze S3 objects by using managed data identifiers that it provides, custom data identifiers that you define, or a combination of the two. You can also fine tune the analysis with allow lists. When you configure settings for automated sensitive data discovery or a sensitive data discovery job, you specify which to use:

- **Managed data identifiers** – These are built-in criteria and techniques that are designed to detect specific types of sensitive data. For example, they can detect credit card numbers, AWS secret access keys, and passport numbers for particular countries and regions. They can detect a large and growing list of sensitive data types for many countries and regions. This includes multiple types of personally identifiable information (PII), financial information, and credentials data. For more information, see Using managed data identifiers.
- **Custom data identifiers** – These are custom criteria that you define to detect sensitive data. Each custom data identifier specifies a regular expression (*regex*) that defines a text pattern to match and, optionally, character sequences and a proximity rule that refine the results. You can use them to detect sensitive data that reflects your particular scenarios, intellectual property,

or proprietary data—for example, employee IDs, customer account numbers, or internal data classifications. For more information, see Building custom data identifiers.

- **Allow lists** – These specify text and text patterns that you want Macie to ignore. You can use them to specify sensitive data exceptions for your particular scenarios or environment—for example, public names or phone numbers for your organization, or sample data that your organization uses for testing. If Macie finds text that matches an entry or pattern in an allow list, Macie doesn't report that occurrence of text. This is the case even if the text matches the criteria of a managed or custom data identifier. For more information, see Defining sensitive data exceptions with allow lists.

When Macie analyzes an S3 object, Macie retrieves the latest version of the object from Amazon S3, and then inspects the object's contents for sensitive data. Macie can analyze an object if the following is true:

- The object uses a supported file or storage format and it's stored in an S3 general purpose bucket using a supported storage class. For more information, see Supported storage classes and formats.
- If the object is encrypted, it's encrypted with a key that Macie can access and is allowed to use. For more information, see Analyzing encrypted S3 objects.
- If the object is stored in a bucket that has a restrictive bucket policy, the policy allows Macie to access objects in the bucket. For more information, see Allowing Macie to access S3 buckets and objects.

To help you meet and maintain compliance with your data security and privacy requirements, Macie produces records of the sensitive data that it finds and the analysis that it performs —*sensitive data findings* and *sensitive data discovery results*. A *sensitive data finding* is a detailed report of sensitive data that Macie found in an S3 object. A *sensitive data discovery result* is a record that logs details about the analysis of an object. Each type of record adheres to a standardized schema, which can help you query, monitor, and process them by using other applications, services, and systems as necessary.

> **Tip**
>
> Although Macie is optimized for Amazon S3, you can use it to discover sensitive data in resources that you currently store elsewhere. You can do this by moving the data to Amazon S3 temporarily or permanently. For example, export Amazon Relational Database

Service or Amazon Aurora snapshots to Amazon S3 in Apache Parquet format. Or export an Amazon DynamoDB table to Amazon S3. You can then create a job to analyze the data in Amazon S3.

Topics

Using managed data identifiers

Amazon Macie uses a combination of criteria and techniques, including machine learning and pattern matching, to detect sensitive data in Amazon Simple Storage Service (Amazon S3) objects. These criteria and techniques, collectively referred to as *managed data identifiers*, can detect a large and growing list of sensitive data types for many countries and regions, including multiple types of credentials data, financial information, personal health information (PHI), and personally identifiable information (PII). Each managed data identifier is designed to detect a specific type of sensitive data—for example, AWS secret access keys, credit card numbers, or passport numbers for a particular country or region.

Macie can detect the following categories of sensitive data by using managed data identifiers:

- Credentials, for credentials data such as private keys and AWS secret access keys.
- Financial information, for financial data such as credit card numbers and bank account numbers.
- Personal information, for PHI such as health insurance and medical identification numbers, and PII such as driver's license identification numbers and passport numbers.

Within each category, Macie can detect multiple types of sensitive data. The topics in this section list and describe each type and any relevant requirements for detecting it. For each type, they also

indicate the unique identifier (ID) for the managed data identifier that's designed to detect the data. When you create a sensitive data discovery job or configure settings for automated sensitive data discovery, you can use these IDs to specify which managed data identifiers you want Macie to use when it analyzes S3 objects.

Topics

For a list of managed data identifiers that we recommend for jobs, see Managed data identifiers recommended for sensitive data discovery jobs. For a list of managed data identifiers that we recommend and are used by default for automated sensitive data discovery, see Default settings for automated sensitive data discovery.

Keyword requirements for managed data identifiers

To detect certain types of sensitive data by using managed data identifiers, Amazon Macie requires a keyword to be in proximity of the data. If this is the case for a particular type of data, reference topics in this section indicate the keyword requirements for that data.

If a keyword has to be in proximity of a particular type of data, the keyword typically has to be within 30 characters (inclusively) of the data. Additional proximity requirements vary based on the file type or storage format of an Amazon Simple Storage Service (Amazon S3) object.

Structured columnar data

For columnar data, a keyword has to be part of the same value or in the name of the column or field that stores a value. This is the case for Microsoft Excel workbooks, CSV files, and TSV files.

For example, if the value for a field contains both *SSN* and a nine-digit number that uses the syntax of a US Social Security number (SSN), Macie can detect the SSN in the field. Similarly, if the name of a column contains *SSN*, Macie can detect each SSN in the column. Macie treats the values in that column as being in proximity of the keyword *SSN*.

Structured record-based data

For record-based data, a keyword has to be part of the same value or in the name of an element in the path to the field or array that stores a value. This is the case for Apache Avro object containers, Apache Parquet files, JSON files, and JSON Lines files.

For example, if the value for a field contains both *credentials* and a character sequence that uses the syntax of an AWS secret access key, Macie can detect the key in the field. Similarly, if the path to a field is `$.credentials.aws.key`, Macie can detect an AWS secret access key in the field. Macie treats the value in the field as being in proximity of the keyword *credentials*.

Unstructured data

For unstructured data, a keyword typically has to be within 30 characters (inclusively) of the data. There aren't any additional proximity requirements. This is the case for Adobe Portable Document Format files, Microsoft Word documents, email messages, and non-binary text files other than CSV, JSON, JSON Lines, and TSV files. This includes any structured data, such as tables or XML, in these types of files.

Keywords aren't case sensitive. In addition, if a keyword contains a space, Macie automatically matches keyword variations that don't contain the space or contain an underscore (_) or a hyphen (-) instead of the space. In certain cases, Macie also expands or abbreviates a keyword to address common variations of the keyword.

For a demonstration of how keywords provide context and help Macie detect specific types of sensitive data, watch the following video: How Amazon Macie uses keywords to discover sensitive data.

Quick reference: Managed data identifiers by type

In Amazon Macie, a *managed data identifier* is a set of built-in criteria and techniques that are designed to detect a specific type of sensitive data—for example, credit card numbers, AWS secret access keys, or passport numbers for a particular country or region. These identifiers can detect a large and growing list of sensitive data types for many countries and regions, including multiple types of credentials data, financial information, personal health information (PHI), and personally identifiable information (PII).

The following table lists all the managed data identifiers that Macie currently provides, organized by sensitive data type. For each type, it provides the following information:

- **Sensitive data category** – Specifies the general category of sensitive data that includes the type: *Credentials*, for credentials data such as private keys; *Financial information*, for financial data such as credit card numbers and bank account numbers; *Personal information: PHI* for personal health information such as health insurance and medical identification numbers; and, *Personal information: PII* for personally identifiable information such as driver's license identification numbers and passport numbers.
- **Managed data identifier ID** – Specifies the unique identifier (ID) for one or more managed data identifiers that are designed to detect the data. When you create a sensitive data discovery job or configure settings for automated sensitive data discovery, you can use these IDs to specify which managed data identifiers you want Macie to use when it analyzes data. For a list of managed data identifiers that we recommend for jobs, see Managed data identifiers recommended for sensitive data discovery jobs. For a list of managed data identifiers that we recommend for automated sensitive data discovery, see Default settings for automated sensitive data discovery.
- **Keyword required** – Specifies whether detection requires a keyword to be in proximity of the data. For information about how Macie uses keywords when it analyzes data, see Keyword requirements.
- **Countries and regions** – Specifies which countries and regions the applicable managed data identifiers are designed for. If the managed data identifiers aren't designed for particular countries and regions, this value is *Any*.

To review additional details about the managed data identifiers for a particular type of sensitive data, choose the type.

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
AWS secret access key	Credentials	AWS_CREDENTIALS	Yes	Any
Bank account number	Financial information	BANK_ACCOUNT_NUMBER (for both Canada and the US)	Yes	Canada, US

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
Basic Bank Account Number (BBAN)	Financial information	Depending on country or region: FRANCE_BANK_ACCOUNT_NUMBER, GERMANY_BANK_ACCOUNT_NUMBER, ITALY_BANK_ACCOUNT_NUMBER, SPAIN_BANK_ACCOUNT_NUMBER, UK_BANK_ACCOUNT_NUMBER	Yes	France, Germany, Italy, Spain, UK
Birth date	Personal information: PII	DATE_OF_BIRTH	Yes	Any
Credit card expiration date	Financial information	CREDIT_CARD_EXPIRATION	Yes	Any
Credit card magnetic stripe data	Financial information	CREDIT_CARD_MAGNETIC_STRIPE	Yes	Any

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
Credit card number	Financial information	CREDIT_CARD_NUMBER (for credit card numbers in proximity of a keyword), CREDIT_CARD_NUMBER_(NO_KEYWORD) (for credit card numbers not in proximity of a keyword)	Varies	Any
Credit card verification code	Financial information	CREDIT_CARD_SECURITY_CODE	Yes	Any

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
Driver's license identification number	Personal information: PII	Depending on country or region: AUSTRALIA_DRIVERS_LICENSE, AUSTRIA_DRIVERS_LICENSE, BELGIUM_DRIVERS_LICENSE, BULGARIA_DRIVERS_LICENSE, CANADA_DRIVERS_LICENSE, CROATIA_DRIVERS_LICENSE, CYPRUS_DRIVERS_LICENSE, CZECHIA_DRIVERS_LICENSE, DENMARK_DRIVERS_LICENSE, DRIVERS_LICENSE (for the US), ESTONIA_DRIVERS_LICENSE,	Yes	Australia, Austria, Belgium, Bulgaria, Canada, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, India, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, UK, US

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		FINLAND_DRIVERS_LICENSE, FRANCE_DRIVERS_LICENSE, GERMANY_DRIVERS_LICENSE, GREECE_DRIVERS_LICENSE, HUNGARY_DRIVERS_LICENSE, INDIA_DRIVERS_LICENSE, IRELAND_DRIVERS_LICENSE, ITALY_DRIVERS_LICENSE, LATVIA_DRIVERS_LICENSE, LITHUANIA_DRIVERS_LICENSE, LUXEMBOURG_DRIVERS_LICENSE, MALTA_DRIVERS_LICENSE, NETHERLANDS_DRIVERS_LICENSE, POLAND_DR		

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		IVERS_LICENSE, PORTUGAL_ DRIVERS_L ICENSE, ROMANIA_D RIVERS_LI CENSE, SLOVAKIA_ DRIVERS_L ICENSE, SLOVENIA_ DRIVERS_L ICENSE, SPAIN_DRI VERS_LICENSE, SWEDEN_DR IVERS_LICENSE, UK_DRIVER S_LICENSE		
Drug Enforcement Agency (DEA) Registration Number	Personal information: PHI	US_DRUG_E NFORCEMEN T_AGENCY_ NUMBER	Yes	US
Electoral roll number	Personal information: PII	UK_ELECTO RAL_ROLL_ NUMBER	Yes	UK
Full name	Personal information: PII	NAME	No	Any, if the name uses a Latin character set

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
Global Positioning System (GPS) coordinates	Personal information: PII	LATITUDE_LONGITUDE	Yes	Any, if the coordinates are in proximity of an English keyword
Google Cloud API key	Credentials	GCP_API_KEY	Yes	Any
Health Insurance Claim Number (HICN)	Personal information: PHI	USA_HEALTH_INSURANCE_CLAIM_NUMBER	Yes	US

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
Health insurance or medical identification number	Personal information: PHI	Depending on country or region: CANADA_HEALTH_NUMBER, EUROPEAN_HEALTH_INSURANCE_CARD_NUMBER, FINLAND_EUROPEAN_HEALTH_INSURANCE_NUMBER, FRANCE_HEALTH_INSURANCE_NUMBER, UK_NHS_NUMBER, USA_MEDICARE_BENEFICIARY_IDENTIFIER	Yes	Canada, EU, Finland, France, UK, US
Healthcare Common Procedure Coding System (HCPCS) code	Personal information: PHI	USA_HEALTHCARE_PROCEDURE_CODE	Yes	US

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
HTTP Basic Authorization header	Credentials	HTTP_BASIC_AUTH_HEADER	No	Any
HTTP cookie	Personal information: PII	HTTP_COOKIE	No	Any

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
International Bank Account Number (IBAN)	Financial information	Depending on country or region: ALBANIA_BANK_ACCOUNT_NUMBER, ANDORRA_BANK_ACCOUNT_NUMBER, BOSNIA_AND_HERZEGOVINA_BANK_ACCOUNT_NUMBER, BRAZIL_BANK_ACCOUNT_NUMBER, BULGARIA_BANK_ACCOUNT_NUMBER, COSTA_RICA_BANK_ACCOUNT_NUMBER, CROATIA_BANK_ACCOUNT_NUMBER, CYPRUS_BANK_ACCOUNT_NUMBER, CZECH_REPUBLIC_BANK_ACCOUNT_NUMBER,	No	Albania, Andorra, Bosnia-Herzegovina, Brazil, Bulgaria, Costa Rica, Croatia, Cyprus, Czech Republic, Denmark, Dominican Republic, Egypt, Estonia, Faroe Islands, Finland, France, Georgia, Germany, Greece, Greenland, Hungary, Iceland, Ireland, Italy, Jordan, Kosovo, Liechtenstein, Lithuania, Malta, Mauritania, Mauritius, Monaco, Montenegro, Netherlands, North Macedonia, Poland, Portugal, San Marino, Senegal, Serbia, Slovakia,

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		DENMARK_BANK_ACCOUNT_NUMBER, DOMINICAN_REPUBLIC_BANK_ACCOUNT_NUMBER, EGYPT_BANK_ACCOUNT_NUMBER, ESTONIA_BANK_ACCOUNT_NUMBER, FAROE_ISLANDS_BANK_ACCOUNT_NUMBER, FINLAND_BANK_ACCOUNT_NUMBER, FRANCE_BANK_ACCOUNT_NUMBER, GEORGIA_BANK_ACCOUNT_NUMBER, GERMANY_BANK_ACCOUNT_NUMBER, GREECE_BANK_ACCOUNT_NUMBER, GREENLAND_BANK_ACC		Slovenia, Spain, Sweden, Switzerland, Timor-Leste, Tunisia, Türkiye, UK, Ukraine, United Arab Emirates, Virgin Islands (British)

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		OUNT_NUMBER, HUNGARY_BANK_ACCOUNT_NUMBER, ICELAND_BANK_ACCOUNT_NUMBER, IRELAND_BANK_ACCOUNT_NUMBER, ITALY_BANK_ACCOUNT_NUMBER, JORDAN_BANK_ACCOUNT_NUMBER, KOSOVO_BANK_ACCOUNT_NUMBER, LIECHTENSTEIN_BANK_ACCOUNT_NUMBER, LITHUANIA_BANK_ACCOUNT_NUMBER, MALTA_BANK_ACCOUNT_NUMBER, MAURITANIA_BANK_ACCOUNT_NUMBER, MAURITIUS_BANK_ACC		

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		OUNT_NUMBER, MONACO_BANK_ACCOUNT_NUMBER, MONTENEGRO_BANK_ACCOUNT_NUMBER, NETHERLANDS_BANK_ACCOUNT_NUMBER, NORTH_MACEDONIA_BANK_ACCOUNT_NUMBER, POLAND_BANK_ACCOUNT_NUMBER, PORTUGAL_BANK_ACCOUNT_NUMBER, SAN_MARINO_BANK_ACCOUNT_NUMBER, SENEGAL_BANK_ACCOUNT_NUMBER, SERBIA_BANK_ACCOUNT_NUMBER, SLOVAKIA_BANK_ACCOUNT_NUMBE		

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		R, SLOVENIA_BANK_ACCOUNT_NUMBER, SPAIN_BANK_ACCOUNT_NUMBER, SWEDEN_BANK_ACCOUNT_NUMBER, SWITZERLAND_BANK_ACCOUNT_NUMBER, TIMOR_LESTE_BANK_ACCOUNT_NUMBER, TUNISIA_BANK_ACCOUNT_NUMBER, TURKIYE_BANK_ACCOUNT_NUMBER, UK_BANK_ACCOUNT_NUMBER, UKRAINE_BANK_ACCOUNT_NUMBER, UNITED_ARAB_EMIRATES_BANK_ACCOUNT_NUMBER, VIRGIN_IS		

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		LANDS_BANK_ACCOUNT_NUMBER (for the British Virgin Islands)		
JSON Web Token (JWT)	Credentials	JSON_WEB_TOKEN	No	Any
Mailing address	Personal information: PII	ADDRESS, BRAZIL_CEP_CODE (for Brazil's Código de Endereçamento Postal)	Varies	Australia, Brazil, Canada, France, Germany, Italy, Spain, UK, US
National Drug Code (NDC)	Personal information: PHI	USA_NATIONAL_DRUG_CODE	Yes	US

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
National identification number	Personal information: PII	Depending on country or region: ARGENTINA_DNI_NUMBER, BRAZIL_RG_NUMBER, CHILE_RUT_NUMBER, COLOMBIA_CITIZENSHIP_CARD_NUMBER, FRANCE_NATIONAL_IDENTIFICATION_NUMBER, GERMANY_NATIONAL_IDENTIFICATION_NUMBER, INDIA_AADHAAR_NUMBER, ITALY_NATIONAL_IDENTIFICATION_NUMBER, MEXICO_CURP_NUMBER, SPAIN_DNI_NUMBER	Yes	Argentina, Brazil, Chile, Colombia, France, Germany, India, Italy, Mexico, Spain

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
National Insurance Number (NINO)	Personal information: PII	UK_NATIONAL_INSURANCE_NUMBER	Yes	UK
National Provider Identifier (NPI)	Personal information: PHI	USA_NATIONAL_PROVIDER_IDENTIFIER	Yes	US
OpenSSH private key	Credentials	OPENSSH_PRIVATE_KEY	No	Any
Passport number	Personal information: PII	Depending on country or region: CANADA_PASSPORT_NUMBER, FRANCE_PASSPORT_NUMBER, GERMANY_PASSPORT_NUMBER, ITALY_PASSPORT_NUMBER, SPAIN_PASSPORT_NUMBER, UK_PASSPORT_NUMBER, USA_PASSPORT_NUMBER	Yes	Canada, France, Germany, Italy, Spain, UK, US

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
Permanent residence number	Personal information: PII	CANADA_NATIONAL_IDENTIFICATION_NUMBER	Yes	Canada
PGP private key	Credentials	PGP_PRIVATE_KEY	No	Any
Phone number	Personal information: PII	Depending on country or region: BRAZIL_PHONE_NUMBER, FRANCE_PHONE_NUMBER, GERMANY_PHONE_NUMBER, ITALY_PHONE_NUMBER, PHONE_NUMBER (for Canada and the US), SPAIN_PHONE_NUMBER, UK_PHONE_NUMBER	Varies	Brazil, Canada, France, Germany, Italy, Spain, UK, US
Public-Key Cryptography Standard (PKCS) private key	Credentials	PKCS	No	Any

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
Public transportation card number	Personal information: PII	ARGENTINA_TARJETA_SUBE	Yes	Argentina
PuTTY private key	Credentials	PUTTY_PRIVATE_KEY	No	Any
Social Insurance Number (SIN)	Personal information: PII	CANADA_SOCIAL_INSURANCE_NUMBER	Yes	Canada
Social Security number (SSN)	Personal information: PII	Depending on country or region: SPAIN_SOCIAL_SECURITY_NUMBER, USA_SOCIAL_SECURITY_NUMBER	Yes	Spain, US
the section called “Stripe API key”	Credentials	STRIPE_CREDENTIALS	No	Any

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
Taxpayer identification or reference number	Personal information: PII	Depending on country or region: ARGENTINA_INDIVIDUAL_TAX_IDENTIFICATION_NUMBER, ARGENTINA_ORGANIZATION_TAX_IDENTIFICATION_NUMBER, AUSTRALIA_TAX_FILE_NUMBER, BRAZIL_CNPJ_NUMBER, BRAZIL_CPF_NUMBER, CHILE_RUT_NUMBER, COLOMBIA_INDIVIDUAL_NIT_NUMBER, COLOMBIA_ORGANIZATION_NIT_NUMBER, FRANCE_TAX_IDENTIFICATION_NUMBER,	Yes	Argentina, Australia, Brazil, Chile, Colombia, France, Germany, India, Italy, Mexico, Spain, UK, US

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		GERMANY_TAX_IDENTIFICATION_NUMBER, INDIA_PERMANENT_ACCOUNT_NUMBER, ITALY_NATIONAL_IDENTIFICATION_NUMBER, MEXICO_INDIVIDUAL_RFC_NUMBER, MEXICO_ORGANIZATION_RFC_NUMBER, SPAIN_NIE_NUMBER, SPAIN_NIF_NUMBER, SPAIN_TAX_IDENTIFICATION_NUMBER, UK_TAX_IDENTIFICATION_NUMBER, USA_INDIVIDUAL_TAX_IDENTIFICATION_NUMBER		

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
Unique device identifier (UDI)	Personal information: PHI	MEDICAL_DEVICE_UDI	Yes	US
Vehicle identification number (VIN)	Personal information: PII	VEHICLE_IDENTIFICATION_NUMBER	Yes	Any, if the VIN is in proximity of a keyword in one of the following languages: English, French, German, Lithuanian, Polish, Portuguese, Romanian, or Spanish

Detailed reference: Managed data identifiers by category

In Amazon Macie, *managed data identifiers* are built-in criteria and techniques that are designed to detect specific types of sensitive data. They can detect a large and growing list of sensitive data types for many countries and regions, including multiple types of credentials data, financial information, and personal information. Each managed data identifier is designed to detect a specific type of sensitive data—for example, AWS secret access keys, credit card numbers, or passport numbers for a particular country or region.

Macie can detect several categories of sensitive data by using managed data identifiers. Within each category, Macie can detect multiple types of sensitive data. The topics in this section list and describe each type and any relevant requirements for detecting the data. You can browse the topics by category:

- Credentials – For credentials data such as private keys and AWS secret access keys.
- Financial information – For financial data such as credit card numbers and bank account numbers.

- Personal information: PHI – For personal health information (PHI) such as health insurance and medical identification numbers.
- Personal information: PII – For personally identifiable information (PII) such as driver's license identification numbers and passport numbers.

Or choose a specific type of sensitive data from the following table. The table lists all the managed data identifiers that Macie currently provides, organized by sensitive data type. The table also summarizes relevant requirements for detecting each type.

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
AWS secret access key	Credentials	AWS_CREDENTIALS	Yes	Any
Bank account number	Financial information	BANK_ACCOUNT_NUMBER (for both Canada and the US)	Yes	Canada, US
Basic Bank Account Number (BBAN)	Financial information	Depending on country or region: FRANCE_BANK_ACCOUNT_NUMBER, GERMANY_BANK_ACCOUNT_NUMBER, ITALY_BANK_ACCOUNT_NUMBER, SPAIN_BANK_ACCOUNT_NUMBER, UK_BANK_A	Yes	France, Germany, Italy, Spain, UK

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		CCOUNT_NUMBER		
Birth date	Personal information: PII	DATE_OF_BIRTH	Yes	Any
Credit card expiration date	Financial information	CREDIT_CARD_EXPIRATION	Yes	Any
Credit card magnetic stripe data	Financial information	CREDIT_CARD_MAGNETIC_STRIPE	Yes	Any
Credit card number	Financial information	CREDIT_CARD_NUMBER (for credit card numbers in proximity of a keyword), CREDIT_CARD_NUMBER_(NO_KEYWORD) (for credit card numbers not in proximity of a keyword)	Varies	Any
Credit card verification code	Financial information	CREDIT_CARD_SECURITY_CODE	Yes	Any

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
Driver's license identification number	Personal information: PII	Depending on country or region: AUSTRALIA_DRIVERS_LICENSE, AUSTRIA_DRIVERS_LICENSE, BELGIUM_DRIVERS_LICENSE, BULGARIA_DRIVERS_LICENSE, CANADA_DRIVERS_LICENSE, CROATIA_DRIVERS_LICENSE, CYPRUS_DRIVERS_LICENSE, CZECHIA_DRIVERS_LICENSE, DENMARK_DRIVERS_LICENSE, DRIVERS_LICENSE (for the US), ESTONIA_DRIVERS_LICENSE,	Yes	Australia, Austria, Belgium, Bulgaria, Canada, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, India, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, UK, US

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		FINLAND_DRIVERS_LICENSE, FRANCE_DRIVERS_LICENSE, GERMANY_DRIVERS_LICENSE, GREECE_DRIVERS_LICENSE, HUNGARY_DRIVERS_LICENSE, INDIA_DRIVERS_LICENSE, IRELAND_DRIVERS_LICENSE, ITALY_DRIVERS_LICENSE, LATVIA_DRIVERS_LICENSE, LITHUANIA_DRIVERS_LICENSE, LUXEMBOURG_DRIVERS_LICENSE, MALTA_DRIVERS_LICENSE, NETHERLANDS_DRIVERS_LICENSE, POLAND_DR		

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		IVERS_LICENSE, PORTUGAL_ DRIVERS_L ICENSE, ROMANIA_D RIVERS_LI CENSE, SLOVAKIA_ DRIVERS_L ICENSE, SLOVENIA_ DRIVERS_L ICENSE, SPAIN_DRI VERS_LICENSE, SWEDEN_DR IVERS_LICENSE, UK_DRIVER S_LICENSE		
Drug Enforcement Agency (DEA) Registration Number	Personal information: PHI	US_DRUG_E NFORCEMEN T_AGENCY_ NUMBER	Yes	US
Electoral roll number	Personal information: PII	UK_ELECTO RAL_ROLL_ NUMBER	Yes	UK
Full name	Personal information: PII	NAME	No	Any, if the name uses a Latin character set

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
Global Positioning System (GPS) coordinates	Personal information: PII	LATITUDE_LONGITUDE	Yes	Any, if the coordinates are in proximity of an English keyword
Google Cloud API key	Credentials	GCP_API_KEY	Yes	Any
Health Insurance Claim Number (HICN)	Personal information: PHI	USA_HEALTH_INSURANCE_CLAIM_NUMBER	Yes	US

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
Health insurance or medical identification number	Personal information: PHI	Depending on country or region: CANADA_HEALTH_NUMBER, EUROPEAN_HEALTH_INSURANCE_CARD_NUMBER, FINLAND_EUROPEAN_HEALTH_INSURANCE_NUMBER, FRANCE_HEALTH_INSURANCE_NUMBER, UK_NHS_NUMBER, USA_MEDICARE_BENEFICIARY_IDENTIFIER	Yes	Canada, EU, Finland, France, UK, US
Healthcare Common Procedure Coding System (HCPCS) code	Personal information: PHI	USA_HEALTHCARE_PROCEDURE_CODE	Yes	US

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
HTTP Basic Authorization header	Credentials	HTTP_BASIC_AUTH_HEADER	No	Any
HTTP cookie	Personal information: PII	HTTP_COOKIE	No	Any

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
International Bank Account Number (IBAN)	Financial information	Depending on country or region: ALBANIA_BANK_ACCOUNT_NUMBER, ANDORRA_BANK_ACCOUNT_NUMBER, BOSNIA_AND_HERZEGOVINA_BANK_ACCOUNT_NUMBER, BRAZIL_BANK_ACCOUNT_NUMBER, BULGARIA_BANK_ACCOUNT_NUMBER, COSTA_RICA_BANK_ACCOUNT_NUMBER, CROATIA_BANK_ACCOUNT_NUMBER, CYPRUS_BANK_ACCOUNT_NUMBER, CZECH_REPUBLIC_BANK_ACCOUNT_NUMBER,	No	Albania, Andorra, Bosnia-Herzegovina, Brazil, Bulgaria, Costa Rica, Croatia, Cyprus, Czech Republic, Denmark, Dominican Republic, Egypt, Estonia, Faroe Islands, Finland, France, Georgia, Germany, Greece, Greenland, Hungary, Iceland, Ireland, Italy, Jordan, Kosovo, Liechtenstein, Lithuania, Malta, Mauritania, Mauritius, Monaco, Montenegro, Netherlands, North Macedonia, Poland, Portugal, San Marino, Senegal, Serbia, Slovakia,

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		DENMARK_BANK_ACCOUNT_NUMBER, DOMINICAN_REPUBLIC_BANK_ACCOUNT_NUMBER, EGYPT_BANK_ACCOUNT_NUMBER, ESTONIA_BANK_ACCOUNT_NUMBER, FAROE_ISLANDS_BANK_ACCOUNT_NUMBER, FINLAND_BANK_ACCOUNT_NUMBER, FRANCE_BANK_ACCOUNT_NUMBER, GEORGIA_BANK_ACCOUNT_NUMBER, GERMANY_BANK_ACCOUNT_NUMBER, GREECE_BANK_ACCOUNT_NUMBER, GREENLAND_BANK_ACC		Slovenia, Spain, Sweden, Switzerland, Timor-Leste, Tunisia, Türkiye, UK, Ukraine, United Arab Emirates, Virgin Islands (British)

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		OUNT_NUMB ER, HUNGARY_B ANK_ACCOU NT_NUMBER , ICELAND_B ANK_ACCOU NT_NUMBER , IRELAND_B ANK_ACCOU NT_NUMBER , ITALY_BAN K_ACCOUNT _NUMBER, JORDAN_BA NK_ACCOUN T_NUMBER, KOSOVO_BA NK_ACCOUN T_NUMBER, LIECHTENS TEIN_BANK _ACCOUNT_ NUMBER, LITHUANIA _BANK_ACC OUNT_NUMB ER, MALTA_BAN K_ACCOUNT _NUMBER, MAURITANI A_BANK_AC COUNT_NUM BER, MAURITIUS _BANK_ACC		

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		OUNT_NUMBER, MONACO_BANK_ACCOUNT_NUMBER, MONTENEGRO_BANK_ACCOUNT_NUMBER, NETHERLANDS_BANK_ACCOUNT_NUMBER, NORTH_MACEDONIA_BANK_ACCOUNT_NUMBER, POLAND_BANK_ACCOUNT_NUMBER, PORTUGAL_BANK_ACCOUNT_NUMBER, SAN_MARINO_BANK_ACCOUNT_NUMBER, SENEGAL_BANK_ACCOUNT_NUMBER, SERBIA_BANK_ACCOUNT_NUMBER, SLOVAKIA_BANK_ACCOUNT_NUMBE		

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		R, SLOVENIA_BANK_ACCOUNT_NUMBER, SPAIN_BANK_ACCOUNT_NUMBER, SWEDEN_BANK_ACCOUNT_NUMBER, SWITZERLAND_BANK_ACCOUNT_NUMBER, TIMOR_LESTE_BANK_ACCOUNT_NUMBER, TUNISIA_BANK_ACCOUNT_NUMBER, TURKIYE_BANK_ACCOUNT_NUMBER, UK_BANK_ACCOUNT_NUMBER, UKRAINE_BANK_ACCOUNT_NUMBER, UNITED_ARAB_EMIRATES_BANK_ACCOUNT_NUMBER, VIRGIN_IS		

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		LANDS_BANK_ACCOUNT_NUMBER (for the British Virgin Islands)		
JSON Web Token (JWT)	Credentials	JSON_WEB_TOKEN	No	Any
Mailing address	Personal information: PII	ADDRESS, BRAZIL_CEP_CODE (for Brazil's Código de Endereçamento Postal)	Varies	Australia, Brazil, Canada, France, Germany, Italy, Spain, UK, US
National Drug Code (NDC)	Personal information: PHI	USA_NATIONAL_DRUG_CODE	Yes	US

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
National identification number	Personal information: PII	Depending on country or region: ARGENTINA_DNI_NUMBER, BRAZIL_RG_NUMBER, CHILE_RUT_NUMBER, COLOMBIA_CITIZENSHIP_CARD_NUMBER, FRANCE_NATIONAL_IDENTIFICATION_NUMBER, GERMANY_NATIONAL_IDENTIFICATION_NUMBER, INDIA_AADHAAR_NUMBER, ITALY_NATIONAL_IDENTIFICATION_NUMBER, MEXICO_CURP_NUMBER, SPAIN_DNI_NUMBER	Yes	Argentina, Brazil, Chile, Colombia, France, Germany, India, Italy, Mexico, Spain

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
National Insurance Number (NINO)	Personal information: PII	UK_NATIONAL_INSURANCE_NUMBER	Yes	UK
National Provider Identifier (NPI)	Personal information: PHI	USA_NATIONAL_PROVIDER_IDENTIFIER	Yes	US
OpenSSH private key	Credentials	OPENSSH_PRIVATE_KEY	No	Any
Passport number	Personal information: PII	Depending on country or region: CANADA_PASSPORT_NUMBER, FRANCE_PASSPORT_NUMBER, GERMANY_PASSPORT_NUMBER, ITALY_PASSPORT_NUMBER, SPAIN_PASSPORT_NUMBER, UK_PASSPORT_NUMBER, USA_PASSPORT_NUMBER	Yes	Canada, France, Germany, Italy, Spain, UK, US

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
Permanent residence number	Personal information: PII	CANADA_NATIONAL_IDENTIFICATION_NUMBER	Yes	Canada
PGP private key	Credentials	PGP_PRIVATE_KEY	No	Any
Phone number	Personal information: PII	Depending on country or region: BRAZIL_PHONE_NUMBER, FRANCE_PHONE_NUMBER, GERMANY_PHONE_NUMBER, ITALY_PHONE_NUMBER, PHONE_NUMBER (for Canada and the US), SPAIN_PHONE_NUMBER, UK_PHONE_NUMBER	Varies	Brazil, Canada, France, Germany, Italy, Spain, UK, US
Public-Key Cryptography Standard (PKCS) private key	Credentials	PKCS	No	Any

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
Public transportation card number	Personal information: PII	ARGENTINA_TARJETA_SUBE	Yes	Argentina
PuTTY private key	Credentials	PUTTY_PRIVATE_KEY	No	Any
Social Insurance Number (SIN)	Personal information: PII	CANADA_SOCIAL_INSURANCE_NUMBER	Yes	Canada
Social Security number (SSN)	Personal information: PII	Depending on country or region: SPAIN_SOCIAL_SECURITY_NUMBER, USA_SOCIAL_SECURITY_NUMBER	Yes	Spain, US
the section called “Stripe API key”	Credentials	STRIPE_CREDENTIALS	No	Any

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
Taxpayer identification or reference number	Personal information: PII	Depending on country or region: ARGENTINA_INDIVIDUAL_TAX_IDENTIFICATION_NUMBER, ARGENTINA_ORGANIZATION_TAX_IDENTIFICATION_NUMBER, AUSTRALIA_TAX_FILE_NUMBER, BRAZIL_CNPJ_NUMBER, BRAZIL_CPF_NUMBER, CHILE_RUT_NUMBER, COLOMBIA_INDIVIDUAL_NIT_NUMBER, COLOMBIA_ORGANIZATION_NIT_NUMBER, FRANCE_TAX_IDENTIFICATION_NUMBER,	Yes	Argentina, Australia, Brazil, Chile, Colombia, France, Germany, India, Italy, Mexico, Spain, UK, US

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
		GERMANY_TAX_IDENTIFICATION_NUMBER, INDIA_PERMANENT_ACCOUNT_NUMBER, ITALY_NATIONAL_IDENTIFICATION_NUMBER, MEXICO_INDIVIDUAL_RFC_NUMBER, MEXICO_ORGANIZATION_RFC_NUMBER, SPAIN_NIE_NUMBER, SPAIN_NIF_NUMBER, SPAIN_TAX_IDENTIFICATION_NUMBER, UK_TAX_IDENTIFICATION_NUMBER, USA_INDIVIDUAL_TAX_IDENTIFICATION_NUMBER		

Sensitive data type	Sensitive data category	Managed data identifier ID	Keyword required	Countries and regions
Unique device identifier (UDI)	Personal information: PHI	MEDICAL_DEVICE_UDI	Yes	US
Vehicle identification number (VIN)	Personal information: PII	VEHICLE_IDENTIFICATION_NUMBER	Yes	Any, if the VIN is in proximity of a keyword in one of the following languages: English, French, German, Lithuanian, Polish, Portuguese, Romanian, or Spanish

Managed data identifiers for credentials data

Amazon Macie can detect multiple types of sensitive credentials data by using managed data identifiers. The topics on this page specify each type and provide information about the managed data identifier that's designed to detect the data. Each topic provides the following information:

- **Managed data identifier ID** – Specifies the unique identifier (ID) for the managed data identifier that's designed to detect the data. When you create a sensitive data discovery job or configure settings for automated sensitive data discovery, you can use this ID to specify whether you want Macie to use the managed data identifier when it analyzes data.
- **Supported countries and regions** – Indicates which countries or regions the applicable managed data identifier is designed for. If the managed data identifier isn't designed for a particular country or region, this value is *Any*.
- **Keyword required** – Specifies whether detection requires a keyword to be in proximity of the data. If a keyword is required, the topic also provides examples of required keywords. For information about how Macie uses keywords when it analyzes data, see Keyword requirements.

- **Comments** – Provides any relevant details that might affect your choice of managed data identifier or your investigation into reported occurrences of the sensitive data. The details include information such as supported standards, syntax requirements, and exceptions.

The topics are listed in alphabetical order by sensitive data type.

Sensitive data types

AWS secret access key

Managed data identifier ID: AWS_CREDENTIALS

Supported countries and regions: Any

Keyword required: Yes. Keywords include: *aws_secret_access_key, credentials, secret access key, secret key, set-awscredential*

Comments: Macie doesn't report occurrences of the following character sequences, which are commonly used as fictitious examples: `je7MtGbClwBF/2Zp9Utk/h3yCo8nvbEXAMPLEKEY` and `wJalrXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY`.

Google Cloud API key

Managed data identifier ID: GCP_API_KEY

Supported countries and regions: Any

Keyword required: Yes. Keywords include: *G_PLACES_KEY, GCP api key, GCP key, google cloud key, google-api-key, google-cloud-apikeys, GOOGLEKEY, X-goog-api-key*

Comments: Macie can detect only the string (`keyString`) component of a Google Cloud API key. Support doesn't include detection of the ID or display name component of a Google Cloud API key.

HTTP Basic Authorization header

Managed data identifier ID: HTTP_BASIC_AUTH_HEADER

Supported countries and regions: Any

Keyword required: No

Comments: Detection requires a complete header, including the field name and authentication scheme directive, as specified by RFC 7617. For example: `Authorization: Basic QWxhZGRpbjpvcGVuIHNlc2FtZQ==` and `Proxy-Authorization: Basic dGVzdDoxMjPCow==`.

JSON Web Token (JWT)

Managed data identifier ID: JSON_WEB_TOKEN

Supported countries and regions: Any

Keyword required: No

Comments: Macie can detect JSON Web Tokens (JWTs) that comply with the requirements specified by RFC 7519 for JSON Web Signature (JWS) structures. The tokens can be signed or unsigned.

OpenSSH private key

Managed data identifier ID: OPENSSH_PRIVATE_KEY

Supported countries and regions: Any

Keyword required: No

Comments: None

PGP private key

Managed data identifier ID: PGP_PRIVATE_KEY

Supported countries and regions: Any

Keyword required: No

Comments: None

Public-Key Cryptography Standard (PKCS) private key

Managed data identifier ID: PKCS

Supported countries and regions: Any

Keyword required: No

Comments: None

PuTTY private key

Managed data identifier ID: PUTTY_PRIVATE_KEY

Supported countries and regions: Any

Keyword required: No

Comments: Macie can detect PuTTY private keys that use the following standard headers and header sequence: `PuTTY-User-Key-File`, `Encryption`, `Comment`, `Public-Lines`, `Private-Lines`, and `Private-MAC`. The header values can contain alphanumeric characters, hyphens (-), and newline characters (\n or \r). `Public-Lines` and `Private-Lines` values can also contain forward slashes (/), plus signs (+), and equal signs (=). `Private-MAC` values can also contain plus signs (+). Support doesn't include detection of private keys with header values that contain other characters, such as spaces or underscores (_). Support also doesn't include detection of private keys that include custom headers.

Stripe API key

Managed data identifier ID: STRIPE_CREDENTIALS

Supported countries and regions: Any

Keyword required: No

Comments: Macie doesn't report occurrences of the following character sequences, which are commonly used in Stripe code examples: `sk_test_4eC39HqLyjWDarjtT1zdp7dc` and `pk_test_TYooMQauvdEDq54NiTphI7jx`.

Managed data identifiers for financial information

Amazon Macie can detect multiple types of sensitive financial information by using managed data identifiers. The topics on this page list each type and provide information about the managed data identifiers that are designed to detect the data. Each topic provides the following information:

- **Managed data identifier ID** – Specifies the unique identifier (ID) for one or more managed data identifiers that are designed to detect the data. When you create a sensitive data discovery job or configure settings for automated sensitive data discovery, you can use these IDs to specify which managed data identifiers you want Macie to use when it analyzes data.
- **Supported countries and regions** – Indicates which countries and regions the applicable managed data identifiers are designed for. If the managed data identifiers aren't designed for particular countries or regions, this value is *Any*.
- **Keyword required** – Specifies whether detection requires a keyword to be in proximity of the data. If a keyword is required, the topic also provides examples of required keywords. For information about how Macie uses keywords when it analyzes data, see Keyword requirements.
- **Comments** – Provides any relevant details that might affect your choice of managed data identifier or your investigation into reported occurrences of the sensitive data. The details include information such as supported standards, syntax requirements, and exceptions.

The topics are listed in alphabetical order by sensitive data type.

Sensitive data types

Bank account number

Macie can detect Canadian and US bank account numbers that consist of 9–17 digit sequences and don't contain any spaces.

Managed data identifier ID: BANK_ACCOUNT_NUMBER

Supported countries and regions: Canada, US

Keyword required: Yes. Keywords include: *bank account, bank acct, checking account, checking acct, deposit account, deposit acct, savings account, savings acct, chequing account, chequing acct*

Comments: This managed data identifier is explicitly designed to detect bank account numbers for Canada and the US. These countries don't use the Basic Bank Account Number (BBAN) or International Bank Account Number (IBAN) formats defined by the ISO international standard for numbering bank accounts, as specified by ISO 13616. To detect bank account numbers for other countries and regions, use the managed data identifiers that are designed for those formats. For more information, see Basic Bank Account Number (BBAN) and International Bank Account Number (IBAN).

Basic Bank Account Number (BBAN)

Macie can detect Basic Bank Account Numbers (BBANs) that conform to the BBAN structure defined by the ISO international standard for numbering bank accounts, as specified by ISO 13616. This includes BBANs that don't contain spaces, or use space or hyphen separators—for example, `NWBK60161331926819`, `NWBK 6016 1331 9268 19`, and `NWBK-6016-1331-9268-19`.

Managed data identifier ID: Depending on country or region, FRANCE_BANK_ACCOUNT_NUMBER, GERMANY_BANK_ACCOUNT_NUMBER, ITALY_BANK_ACCOUNT_NUMBER, SPAIN_BANK_ACCOUNT_NUMBER, UK_BANK_ACCOUNT_NUMBER

Supported countries and regions: France, Germany, Italy, Spain, UK

Keyword required: Yes. The following table lists the keywords that Macie recognizes for specific countries and regions.

Country or region	Keywords
France	account code, account number, accountno#, accountnumber#, bban, code bancaire, compte bancaire, customer account id, customer account number, customer bank account id, iban, numéro de compte

Country or region	Keywords
Germany	account code, account number, accountno #, accountnumber#, bankleitzahl, bban, customer account id, customer account number, customer bank account id, geheimzahl, iban, kartennummer, kontonummer, kreditkartennummer, sepa
Italy	account code, account number, accountno #, accountnumber#, bban, codice bancario, conto bancario, customer account id, customer account number, customer bank account id, iban, numero di conto
Spain	account code, account number, accountno #, accountnumber#, bban, código cuenta, código cuenta bancaria, cuenta cliente id, customer account ID, customer account number, customer bank account id, iban, número cuenta bancaria cliente, número cuenta cliente
UK	account code, account number, accountno#, accountnumber#, bban, customer account id, customer account number, customer bank account id, iban, sepa

Comments: These managed data identifiers can also detect International Bank Account Numbers (IBANs) that comply with the ISO 13616 standard. For more information, see International Bank Account Number (IBAN). The managed data identifier for the UK (UK_BANK_ACCOUNT_NUMBER) can also detect domestic bank account numbers for the UK—for example, `60-16-13 31926819`.

Credit card expiration date

Managed data identifier ID: CREDIT_CARD_EXPIRATION

Supported countries and regions: Any

Keyword required: Yes. Keywords include: *exp d, exp m, exp y, expiration, expiry*

Comments: Support includes most date formats, such as all digits and combinations of digits and names of months. Date components can be separated by slashes (/), hyphens (-), or applicable keywords. For example, Macie can detect dates such as `02/26`, `02/2026`, `Feb 2026`, `26-Feb`, and `expY=2026, expM=02`.

Credit card magnetic stripe data

Managed data identifier ID: CREDIT_CARD_MAGNETIC_STRIPE

Supported countries and regions: Any

Keyword required: Yes. Keywords include: *card data, iso7813, mag, magstripe, stripe, swipe*

Comments: Support includes tracks 1 and 2.

Credit card number

Managed data identifier ID: CREDIT_CARD_NUMBER for credit card numbers that are in proximity of a keyword, CREDIT_CARD_NUMBER_(NO_KEYWORD) for credit card numbers that aren't in proximity of a keyword

Supported countries and regions: Any

Keyword required: Varies. Keywords are required by the CREDIT_CARD_NUMBER managed data identifier. Keywords include: *account number, american express, amex, bank card, c card, card, cc #, ccn, check card, cred card, credit, credit card, credit cards, credit no, credit num, dankort, debit, debit card, debit no, debit num, diners club, discover, electron, japanese card bureau, jcb, mastercard, mc, pan, payment account number, payment card number, pcn, pmnt #, pmnt card, pmnt no, pmnt number, union pay, visa*. Keywords aren't required by the CREDIT_CARD_NUMBER_(NO_KEYWORD) managed data identifier.

Comments: Detection requires the data to be a 13–19 digit sequence that adheres to the Luhn check formula and uses a standard card number prefix for any of the following types of credit cards: American Express, Dankort, Diner's Club, Discover, Electron, Japanese Card Bureau (JCB), Mastercard, UnionPay, and Visa.

Macie doesn't report occurrences of the following sequences, which credit card issuers have reserved for public testing: `122000000000003`, `2222405343248877`, `2222990905257051`, `2223007648726984`, `2223577120017656`, `30569309025904`, `34343434343434`,

3528000700000000, 3530111333300000, 3566002020360505, 36148900647913, 36700102000000, 371449635398431, 378282246310005, 378734493671000, 38520000023237, 4012888888881881, 4111111111111111, 4222222222222, 4444333322221111, 4462030000000000, 4484070000000000, 4911830000000, 4917300800000000, 4917610000000000, 4917610000000000003, 5019717010103742, 5105105105105100, 5111010030175156, 5185540810000019, 5200828282828210, 5204230080000017, 5204740009900014, 5420923878724339, 5454545454545454, 5455330760000018, 5506900490000436, 5506900490000444, 5506900510000234, 5506920809243667, 5506922400634930, 5506927427317625, 5553042241984105, 5555553753048194, 5555555555554444, 5610591081018250, 6011000990139424, 6011000400000000, 6011111111111117, 63049017740292441, 63049506000000000, 6331101999990016, 6759649826438453, 6799990100000000019, and 76009244561.

Credit card verification code

Managed data identifier ID: CREDIT_CARD_SECURITY_CODE

Supported countries and regions: Any

Keyword required: Yes. Keywords include: *card id, card identification code, card identification number, card security code, card validation code, card validation number, card verification data, card verification value, cvc, cvc2, cvv, cvv2, elo verification code*

Comments: None

International Bank Account Number (IBAN)

Macie can detect International Bank Account Numbers (IBANs) that consist of up to 34 alphanumeric characters, including elements such as country code. More specifically, Macie can detect IBANs that comply with the ISO international standard for numbering bank accounts, as specified by ISO 13616. This includes IBANs that don't contain spaces, or use space or hyphen separators—for example, GB29NWBK60161331926819, GB29 NWBK 6016 1331 9268 19, and GB29-NWBK-6016-1331-9268-19. Detection includes validation checks based on the Modulus 97 scheme.

Managed data identifier ID: Depending on country or region, ALBANIA_BANK_ACCOUNT_NUMBER, ANDORRA_BANK_ACCOUNT_NUMBER, BOSNIA_AND_HERZEGOVINA_BANK_ACCOUNT_NUMBER, BRAZIL_BANK_ACCOUNT_NUMBER, BULGARIA_BANK_ACCOUNT_NUMBER, COSTA_RICA_BANK_ACCOUNT_NUMBER, CROATIA_BANK_ACCOUNT_NUMBER, CYPRUS_BANK_ACCOUNT_NUMBER,

CZECH_REPUBLIC_BANK_ACCOUNT_NUMBER, DENMARK_BANK_ACCOUNT_NUMBER, DOMINICAN_REPUBLIC_BANK_ACCOUNT_NUMBER, EGYPT_BANK_ACCOUNT_NUMBER, ESTONIA_BANK_ACCOUNT_NUMBER, FAROE_ISLANDS_BANK_ACCOUNT_NUMBER, FINLAND_BANK_ACCOUNT_NUMBER, FRANCE_BANK_ACCOUNT_NUMBER, GEORGIA_BANK_ACCOUNT_NUMBER, GERMANY_BANK_ACCOUNT_NUMBER, GREECE_BANK_ACCOUNT_NUMBER, GREENLAND_BANK_ACCOUNT_NUMBER, HUNGARY_BANK_ACCOUNT_NUMBER, ICELAND_BANK_ACCOUNT_NUMBER, IRELAND_BANK_ACCOUNT_NUMBER, ITALY_BANK_ACCOUNT_NUMBER, JORDAN_BANK_ACCOUNT_NUMBER, KOSOVO_BANK_ACCOUNT_NUMBER, LIECHTENSTEIN_BANK_ACCOUNT_NUMBER, LITHUANIA_BANK_ACCOUNT_NUMBER, MALTA_BANK_ACCOUNT_NUMBER, MAURITANIA_BANK_ACCOUNT_NUMBER, MAURITIUS_BANK_ACCOUNT_NUMBER, MONACO_BANK_ACCOUNT_NUMBER, MONTENEGRO_BANK_ACCOUNT_NUMBER, NETHERLANDS_BANK_ACCOUNT_NUMBER, NORTH_MACEDONIA_BANK_ACCOUNT_NUMBER, POLAND_BANK_ACCOUNT_NUMBER, PORTUGAL_BANK_ACCOUNT_NUMBER, SAN_MARINO_BANK_ACCOUNT_NUMBER, SENEGAL_BANK_ACCOUNT_NUMBER, SERBIA_BANK_ACCOUNT_NUMBER, SLOVAKIA_BANK_ACCOUNT_NUMBER, SLOVENIA_BANK_ACCOUNT_NUMBER, SPAIN_BANK_ACCOUNT_NUMBER, SWEDEN_BANK_ACCOUNT_NUMBER, SWITZERLAND_BANK_ACCOUNT_NUMBER, TIMOR_LESTE_BANK_ACCOUNT_NUMBER, TUNISIA_BANK_ACCOUNT_NUMBER, TURKIYE_BANK_ACCOUNT_NUMBER, UK_BANK_ACCOUNT_NUMBER, UKRAINE_BANK_ACCOUNT_NUMBER, UNITED_ARAB_EMIRATES_BANK_ACCOUNT_NUMBER, VIRGIN_ISLANDS_BANK_ACCOUNT_NUMBER (for the British Virgin Islands)

Supported countries and regions: Albania, Andorra, Bosnia-Herzegovina, Brazil, Bulgaria, Costa Rica, Croatia, Cyprus, Czech Republic, Denmark, Dominican Republic, Egypt, Estonia, Faroe Islands, Finland, France, Georgia, Germany, Greece, Greenland, Hungary, Iceland, Ireland, Italy, Jordan, Kosovo, Liechtenstein, Lithuania, Malta, Mauritania, Mauritius, Monaco, Montenegro, Netherlands, North Macedonia, Poland, Portugal, San Marino, Senegal, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Timor-Leste, Tunisia, Türkiye, UK, Ukraine, United Arab Emirates, Virgin Islands (British)

Keyword required: No

Comments: The managed data identifiers for France, Germany, Italy, Spain, and the UK can also detect Basic Bank Account Numbers (BBANs) that conform to the BBAN structure defined by the ISO 13616 standard, if the character sequence is in proximity of a keyword. For more information, see Basic Bank Account Number (BBAN).

Managed data identifiers for PHI

Amazon Macie can detect multiple types of sensitive, personal health information (PHI) by using managed data identifiers. The topics on this page specify each type and provide information about the managed data identifier that's designed to detect the data. Each topic provides the following information:

- **Managed data identifier ID** – Specifies the unique identifier (ID) for the managed data identifier that's designed to detect the data. When you create a sensitive data discovery job or configure settings for automated sensitive data discovery, you can use this ID to specify whether you want Macie to use the managed data identifier when it analyzes data.
- **Supported countries and regions** – Indicates which countries or regions the applicable managed data identifier is designed for. If the managed data identifier isn't designed for a particular country or region, this value is *Any*.
- **Keyword required** – Specifies whether detection requires a keyword to be in proximity of the data. If a keyword is required, the topic also provides examples of required keywords. For information about how Macie uses keywords when it analyzes data, see Keyword requirements.
- **Comments** – Provides any relevant details that might affect your choice of managed data identifier or your investigation into reported occurrences of the sensitive data. The details include information such as supported standards, syntax requirements, and exceptions.

The topics are listed in alphabetical order by sensitive data type.

Sensitive data types

Drug Enforcement Agency (DEA) Registration Number

Managed data identifier ID: US_DRUG_ENFORCEMENT_AGENCY_NUMBER

Supported countries and regions: US

Keyword required: Yes. Keywords include: *dea number, dea registration*

Comments: None

Health Insurance Claim Number (HICN)

Managed data identifier ID: USA_HEALTH_INSURANCE_CLAIM_NUMBER

Supported countries and regions: US

Keyword required: Yes. Keywords include: *health insurance claim number, hic no, hic no., hic number, hic#, hicn, hicn#., hicno#*

Comments: None

Health insurance or medical identification number

Support includes European Health Insurance Card numbers for the EU and Finland, health insurance numbers for France, Medicare Beneficiary Identifiers for the US, NHS numbers for the UK, and Personal Health Numbers for Canada.

Managed data identifier ID: Depending on country or region, CANADA_HEALTH_NUMBER, EUROPEAN_HEALTH_INSURANCE_CARD_NUMBER, FINLAND_EUROPEAN_HEALTH_INSURANCE_NUMBER, FRANCE_HEALTH_INSURANCE_NUMBER, UK_NHS_NUMBER, USA_MEDICARE_BENEFICIARY_IDENTIFIER

Supported countries and regions: Canada, EU, Finland, France, UK, US

Keyword required: Yes. The following table lists the keywords that Macie recognizes for specific countries and regions.

Country or region	Keywords
Canada	canada healthcare number, msp number, personal healthcare number, phn, soins de santé
EU	assicurazione sanitaria numero, carta assicurazione numero, carte d’assurance maladie, carte européenne d'assurance maladie, ceam,

Country or region	Keywords
	ehic, ehic#, finlandehicnumber#, gesundheitskarte, hälsokort, health card, health card number, health insurance card, health insurance number, insurance card number, krankenversicherungskarte, krankenversicherungsnummer, medical account number, numero conto medico, numéro d'assurance maladie, numéro de carte d'assurance, numéro de compte medical, número de cuenta médica, número de seguro de salud, número de tarjeta de seguro, sairaanhoitokortin, sairausvakuutuskortti, sairausvakuutusnumero, sjukförsäkring nummer, sjukförsäkringskort, suomi ehic-numero, tarjeta de salud, terveyskortti, tessera sanitaria assicurazione numero, versicherungsnummer
Finland	ehic, ehic#, finland health insurance card, finlandehicnumber#, finska sjukförsäkringskort, hälsokort, health card, health card number, health insurance card, health insurance number, sairaanhoitokortin, sairaanhoitokortin, sairausvakuutuskortti, sairausvakuutusnumero, sjukförsäkring nummer, sjukförsäkringskort, suomen sairausvakuutuskortti, suomi ehic-numero, terveyskortti
France	carte d'assuré social, carte vitale, insurance card
UK	national health service, NHS
US	mbi, medicare beneficiary

Comments: None

Healthcare Common Procedure Coding System (HCPCS) code

Managed data identifier ID: USA_HEALTHCARE_PROCEDURE_CODE

Supported countries and regions: US

Keyword required: Yes. Keywords include: *current procedural terminology, hcpcs, healthcare common procedure coding system*

Comments: None

National Drug Code (NDC)

Managed data identifier ID: USA_NATIONAL_DRUG_CODE

Supported countries and regions: US

Keyword required: Yes. Keywords include: *national drug code, ndc*

Comments: None

National Provider Identifier (NPI)

Managed data identifier ID: USA_NATIONAL_PROVIDER_IDENTIFIER

Supported countries and regions: US

Keyword required: Yes. Keywords include: *hipaa, n.p.i, national provider, npi*

Comments: None

Unique device identifier (UDI)

Managed data identifier ID: MEDICAL_DEVICE_UDI

Supported countries and regions: US

Keyword required: Yes. Keywords include: *blood, blood bag, dev id, device id, device identifier, gs1, hibcc, iccbba, med, udi, unique device id, unique device identifier*

Comments: Macie can detect unique device identifiers (UDIs) that comply with formats approved by the US Food and Drug Administration. This includes standard formats defined by GS1, HIBCC, and ICCBBA. ICCBA support is for the ISBT standard.

Managed data identifiers for PII

Amazon Macie can detect multiple types of sensitive, personally identifiable information (PII) by using managed data identifiers. The topics on this page list each type and provide information about the managed data identifiers that are designed to detect the data. Each topic provides the following information:

- **Managed data identifier ID** – Specifies the unique identifier (ID) for one or more managed data identifiers that are designed to detect the data. When you create a sensitive data discovery job or configure settings for automated sensitive data discovery, you can use these IDs to specify which managed data identifiers you want Macie to use when it analyzes data.
- **Supported countries and regions** – Indicates which countries and regions the applicable managed data identifiers are designed for. If the managed data identifiers aren't designed for particular countries or regions, this value is *Any*.
- **Keyword required** – Specifies whether detection requires a keyword to be in proximity of the data. If a keyword is required, the topic also provides examples of required keywords. For information about how Macie uses keywords when it analyzes data, see Keyword requirements.
- **Comments** – Provides any relevant details that might affect your choice of managed data identifier or your investigation into reported occurrences of the sensitive data. The details include information such as supported standards, syntax requirements, and exceptions.

The topics are listed in alphabetical order by sensitive data type.

Sensitive data types

- Birth date
- Driver's license identification number
- Electoral roll number
- Full name
- Global Positioning System (GPS) coordinates
- HTTP cookie
- Mailing address
- National identification number
- National Insurance Number (NINO)
- Passport number

- Permanent residence number
- Phone number
- Public transportation card number
- Social Insurance Number (SIN)
- Social Security number (SSN)
- Taxpayer identification or reference number
- Vehicle identification number (VIN)

Birth date

Managed data identifier ID: DATE_OF_BIRTH

Supported countries and regions: Any

Keyword required: Yes. Keywords include: *bday, b-day, birth date, birthday, date of birth, dob*

Comments: Support includes most date formats, such as all digits and combinations of digits and names of months. Date components can be separated by spaces, slashes (/), or hyphens (-).

Driver's license identification number

Managed data identifier ID: Depending on country or region, AUSTRALIA_DRIVERS_LICENSE, AUSTRIA_DRIVERS_LICENSE, BELGIUM_DRIVERS_LICENSE, BULGARIA_DRIVERS_LICENSE, CANADA_DRIVERS_LICENSE, CROATIA_DRIVERS_LICENSE, CYPRUS_DRIVERS_LICENSE, CZECHIA_DRIVERS_LICENSE, DENMARK_DRIVERS_LICENSE, DRIVERS_LICENSE (for the US), ESTONIA_DRIVERS_LICENSE, FINLAND_DRIVERS_LICENSE, FRANCE_DRIVERS_LICENSE, GERMANY_DRIVERS_LICENSE, GREECE_DRIVERS_LICENSE, HUNGARY_DRIVERS_LICENSE, INDIA_DRIVERS_LICENSE, IRELAND_DRIVERS_LICENSE, ITALY_DRIVERS_LICENSE, LATVIA_DRIVERS_LICENSE, LITHUANIA_DRIVERS_LICENSE, LUXEMBOURG_DRIVERS_LICENSE, MALTA_DRIVERS_LICENSE, NETHERLANDS_DRIVERS_LICENSE, POLAND_DRIVERS_LICENSE, PORTUGAL_DRIVERS_LICENSE, ROMANIA_DRIVERS_LICENSE, SLOVAKIA_DRIVERS_LICENSE, SLOVENIA_DRIVERS_LICENSE, SPAIN_DRIVERS_LICENSE, SWEDEN_DRIVERS_LICENSE, UK_DRIVERS_LICENSE

Supported countries and regions: Australia, Austria, Belgium, Bulgaria, Canada, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, India, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, UK, US

Keyword required: Yes. The following table lists the keywords that Macie recognizes for specific countries and regions.

Country or region	Keywords
Australia	dl#, dl:, dlno#, driver licence, driver license, driver permit, drivers lic., drivers licence, driver's licence, drivers license, driver's license, drivers permit, driver's permit, drivers permit number, driving licence, driving license, driving permit
Austria	führerschein, fuhrerschein, führerschein republik österreich, fuhrerschein republik osterreich
Belgium	fuehrerschein, fuehrerschein- nr, fuehrersc heinnummer, fuhrerschein, führerschein, fuhrerschein- nr, führerschein- nr, fuhrersch einnummer, führerscheinnummer, numéro permis conduire, permis de conduire, rijbewijs, rijbewijsnummer
Bulgaria	превозно средство, свидетелство за управление на моторно, свидетелство за управление на мпс, сумпс, шофьорска книжка
Canada	dl#, dl:, dlno#, driver licence, driver licences, driver license, driver licenses, driver permit, drivers lic., drivers licence, driver's licence, drivers licences, driver's licences, drivers license, driver's license, drivers licenses, driver's licenses, drivers permit, driver's permit, drivers permit number, driving licence, driving license, driving permit, permis de conduire
Croatia	vozačka dozvola

Country or region	Keywords
Cyprus	άδεια οδήγησης
Czech Republic	číslo licence, číslo licence řidiče, číslo řidičskéh o průkazu, ovladače lic., povolení k jízdě, povolení řidiče, řidiči povolení, řidičský prúkaz, řidičský průkaz
Denmark	kørekort, kørekortnummer
Estonia	juhi litsentsi number, juhiloa number, juhiluba, juhiluba number
Finland	ajokortin numero, ajokortti, förare lic., körkort, körkort nummer, kuljettaja lic., permis de conduire
France	permis de conduire
Germany	fuehrerschein, fuehrerschein- nr, fuehrersc heinnummer, fuhrerschein, führerschein, fuhrerschein- nr, führerschein- nr, fuhrersch einnummer, führerscheinnummer
Greece	δεια οδήγησης, adeia odigisis
Hungary	illesztőprogramok lic, jogosítvány, jogsi, licencszám, vezető engedély, vezetői engedély
India	driver licence, driver licences, driver license, driver licenses, drivers lic., drivers licence, driver's licence, drivers licences, driver's licences, drivers license, driver's license, drivers licenses, driver's licenses, driving licence, driving license
Ireland	ceadúnas tiomána

Country or region	Keywords
Italy	patente di guida, patente di guida numero, patente guida, patente guida numero
Latvia	autovadītāja apliecība, licences numurs, vadītāja apliecība, vadītāja apliecības numurs, vadītāja atļauja, vadītāja licences numurs, vadītāji lic.
Lithuania	vairuotojo pažymėjimas
Luxembourg	fahrerlaubnis, führerschäin
Malta	liċenzja tas-sewqan
Netherlands	permis de conduire, rijbewijs, rijbewijsnummer
Poland	numer licencyjny, prawo jazdy, zezwolenie na prowadzenie
Portugal	carta de condução, carteira de habilitação, carteira de motorist, carteira habilitação, carteira motorist, licença condução, licença de condução, número de licença, número licença, permissão condução, permissão de condução
Romania	numărul permisului de conducere, permis de conducere
Slovakia	číslo licencie, číslo vodičského preukazu, ovládače lic., povolenia vodičov, povolenie jazdu, povolenie na jazdu, povolenie vodiča, vodičský preukaz
Slovenia	vozniško dovoljenje

Country or region	Keywords
Spain	carnet conducer, el carnet de conducer, licencia conducer, licencia de manejo, número carnet conducer, número de carnet de conducer, número de permiso conducer, número de permiso de conducer, número licencia conducer, número permiso conducer, permiso conducción, permiso conducer, permiso de conducción
Sweden	ajokortin numero, dlno# ajokortti, drivere lic., förare lic., körkort, körkort nummer, körkortsn ummer, kuljettajat lic.
UK	dl#, dl:, dlno#, driver licence, driver licences, driver license, driver licenses, driver permit, drivers lic., drivers licence, driver's licence, drivers licences, driver's licences, drivers license, driver's license, drivers licenses, driver's licenses, drivers permit, driver's permit, drivers permit number, driving licence, driving license, driving permit
US	dl#, dl:, dlno#, driver licence, driver licences, driver license, driver licenses, driver permit, drivers lic., drivers licence, driver's licence, drivers licences, driver's licences, drivers license, driver's license, drivers licenses, driver's licenses, drivers permit, driver's permit, drivers permit number, driving licence, driving license, driving permit

Comments: None

Electoral roll number

Managed data identifier ID: UK_ELECTORAL_ROLL_NUMBER

Supported countries and regions: UK

Keyword required: Yes. Keywords include: *electoral #, electoral number, electoral roll #, electoral roll no., electoral roll number, electoralrollno*

Comments: None

Full name

Managed data identifier ID: NAME

Supported countries and regions: Any

Keyword required: No

Comments: Macie can detect full names only. Support is limited to Latin character sets.

Global Positioning System (GPS) coordinates

Managed data identifier ID: LATITUDE_LONGITUDE

Supported countries and regions: Any, if the coordinates are in proximity of an English keyword.

Keyword required: Yes. Keywords include: *coordinate, coordinates, lat long, latitude longitude, position*

Comments: Macie can detect GPS coordinates if the latitude and longitude coordinates are stored as a pair and they're in Decimal Degrees (DD) format, for example `41.948614,-87.655311`. Support doesn't include detection of coordinates in: Degrees Decimal Minutes (DDM) format, for example `41°56.9168'N 87°39.3187'W`; or Degrees, Minutes, Seconds (DMS) format, for example `41°56'55.0104"N 87°39'19.1196"W`.

HTTP cookie

Managed data identifier ID: HTTP_COOKIE

Supported countries and regions: Any

Keyword required: No

Comments: Detection requires a complete `Cookie` or `Set-Cookie` header. The header can include one or more name-value pairs, for example: `Set-Cookie: id=TWlrZQ` and `Cookie: session=3948; lang=en`.

Mailing address

Managed data identifier ID: ADDRESS (for Australia, Canada, France, Germany, Italy, Spain, UK, and the US), BRAZIL_CEP_CODE (for Brazil's Código de Endereçamento Postal)

Supported countries and regions: Australia, Brazil, Canada, France, Germany, Italy, Spain, UK, US

Keyword required: Varies. Keywords aren't required by the ADDRESS managed data identifier. Keywords are required by the BRAZIL_CEP_CODE managed data identifier. Keywords include: *cep, código de endereçamento postal, codigo de endereçamento postal, código postal, codigo postal*

Comments: Although a keyword isn't required by the ADDRESS managed data identifier, detection requires an address to include the name of a city or place and a corresponding ZIP or Postal Code in a supported country or region. The BRAZIL_CEP_CODE managed data identifier can detect only the Código de Endereçamento Postal (CEP) portion of an address.

National identification number

Support includes: Aadhaar numbers for India; Cédula de Ciudadanía numbers for Colombia; Clave Única de Registro de Población (CURP) numbers for Mexico; Codice Fiscale numbers for Italy; Documento Nacional de Identidad (DNI) numbers for Argentina and Spain; French National Institute for Statistics and Economic Studies (INSEE) codes; German National Identity Card numbers; Registro Geral (RG) numbers for Brazil; and, Rol Único Nacional (RUN) numbers for Chile.

Managed data identifier ID: Depending on country or region, ARGENTINA_DNI_NUMBER, BRAZIL_RG_NUMBER, CHILE_RUT_NUMBER, COLOMBIA_CITIZENSHIP_CARD_NUMBER, FRANCE_NATIONAL_IDENTIFICATION_NUMBER, GERMANY_NATIONAL_IDENTIFICATION_NUMBER, INDIA_AADHAAR_NUMBER, ITALY_NATIONAL_IDENTIFICATION_NUMBER, MEXICO_CURP_NUMBER, SPAIN_DNI_NUMBER

Supported countries and regions: Argentina, Brazil, Chile, Colombia, France, Germany, India, Italy, Mexico, Spain

Keyword required: Yes. The following table lists the keywords that Macie recognizes for specific countries and regions.

Country or region	Keywords
Argentina	dni, dni#, d.n.i., documento nacional de identidad

Country or region	Keywords
Brazil	registro geral, rg
Chile	identidad número, nacional identidad, national unique role, nationaluniqueroleID#, número identificación, rol único nacional, rol único tributario, run, run#, r.u.n., rut, rut#, r.u.t., unique national number, unique national role, unique tax registry, unique tax role, unique tributary number, unique tributary role
Colombia	cédula de ciudadanía, documento de identific ación
France	assurance sociale, carte nationale d’identit é, cni, code sécurité sociale, French social security number, fssn#, insee, insurance number, national id number, nationalid#, numéro d'assurance, sécurité sociale, sécurité sociale non., sécurité sociale numéro, social, social security, social security number, socialsecuritynumber, ss#, ssn, ssn#
Germany	ausweisnummer, id number, identification number, identity number, insurance number, personal id, personalausweis
India	aadhaar, aadhar, adhaar, uidai
Italy	codice fiscal, dati anagrafici, ehic, health card, health insurance card, p. iva, partita i.v.a., personal data, tax code, tessera sanitaria

Country or region	Keywords
Mexico	clave personal identidad, clave única, clave única de registro de población, clavepers onalIdentidad, curp, registration code, registry code, personal identidad clave, population code
Spain	dni, dni#, dninúmero#, documento nacional de identidad, identidad único, identidad único#, insurance number, national identific ation number, national identity, nationalid#, nationalidno#, número nacional identidad , personal identification number, personal identity no, unique identity number, uniqueid#

Comments: The managed data identifier for Chile (CHILE_RUT_NUMBER) is designed to detect both Rol Único Nacional (RUN) numbers and Rol Único Tributario (RUT) numbers. For either type of number, Macie doesn't report occurrences where all the digits are zeroes, such as `00000000-K`, because they're commonly used as examples.

Although DNI numbers for Argentina and Spain have different syntaxes, there are similarities between them. Therefore, Macie might report a DNI number for Argentina as a DNI number for Spain, or the other way around. In addition, Macie doesn't report occurrences of the following character sequences, which are commonly used as example DNI numbers: 99999999 and `99.999.999`. Macie also doesn't report occurrences that consist of only zeroes—for example, `000000000` and `00.000.000`.

National Insurance Number (NINO)

Managed data identifier ID: UK_NATIONAL_INSURANCE_NUMBER

Supported countries and regions: UK

Keyword required: Yes. Keywords include: *insurance no., insurance number, insurance#, national insurance number, nationalinsurance#, nationalinsurancenumber, nin, nino*

Comments: None

Passport number

Managed data identifier ID: Depending on country or region, CANADA_PASSPORT_NUMBER, FRANCE_PASSPORT_NUMBER, GERMANY_PASSPORT_NUMBER, ITALY_PASSPORT_NUMBER, SPAIN_PASSPORT_NUMBER, UK_PASSPORT_NUMBER, USA_PASSPORT_NUMBER

Supported countries and regions: Canada, France, Germany, Italy, Spain, UK, US

Keyword required: Yes. The following table lists the keywords that Macie recognizes for specific countries and regions.

Country or region	Keywords
Canada	passeport, passeport#, passport, passport#, passportno, passportno#
France	numéro de passeport, passeport, passeport #, passeport n °, passeport non
Germany	ausstellungsdatum, ausstellungsort, geburtsdatum, passport, passports, reisepass, reisepass–nr, reisepassnummer
Italy	italian passport number, numéro passeport , numéro passeport italien, passaporto, passaporto italiana, passaporto numero, passport number, repubblica italiana passaporto
Spain	españa pasaporte, libreta pasaporte, número pasaporte, pasaporte, passport, passport book, passport no, passport number, spain passport
UK	passeport #, passeport n °, passeport non, passeportn °, passport #, passport no, passport number, passport#, passportid
US	passport, travel document

Comments: None

Permanent residence number

Managed data identifier ID: CANADA_NATIONAL_IDENTIFICATION_NUMBER

Supported countries and regions: Canada

Keyword required: Yes. Keywords include: *carte résident permanent, numéro carte résident permanent, numéro résident permanent, permanent resident card, permanent resident card number, permanent resident no, permanent resident no., permanent resident number, pr no, pr no., pr non, pr number, résident permanent no., résident permanent non*

Comments: None

Phone number

Managed data identifier ID: Depending on country or region, BRAZIL_PHONE_NUMBER, FRANCE_PHONE_NUMBER, GERMANY_PHONE_NUMBER, ITALY_PHONE_NUMBER, PHONE_NUMBER (for Canada and the US), SPAIN_PHONE_NUMBER, UK_PHONE_NUMBER

Supported countries and regions: Brazil, Canada, France, Germany, Italy, Spain, UK, US

Keyword required: Varies. If a keyword is in proximity of the data, the number doesn't have to include a country code. Keywords include: *cell, contact, fax, fax number, mobile, phone, phone number, tel, telephone, telephone number*. For Brazil, keywords also include: *cel, celular, fone, móvel, número residencial, numero residencial, telefone*. If a keyword isn't in proximity of the data, the number has to include a country code.

Comments: For the US, support includes toll-free numbers.

Public transportation card number

Managed data identifier ID: ARGENTINA_TARJETA_SUBE

Supported countries and regions: Argentina

Keyword required: Yes. Keywords include: *sistema único de boleto electrónico, sube*

Comments: Macie can detect 16-digit Sistema Único de Boleto Electrónico (SUBE) card numbers that begin with 6061 and adhere to the Luhn check formula. Card number components can be

separated by spaces or hyphens (-), or not use a separator—for example, `6061 1234 1234 1234`, `6061-1234-1234-1234`, and `6061123412341234`.

Social Insurance Number (SIN)

Managed data identifier ID: CANADA_SOCIAL_INSURANCE_NUMBER

Supported countries and regions: Canada

Keyword required: Yes. Keywords include: *canadian id, numéro d'assurance sociale, sin, social insurance number*

Comments: None

Social Security number (SSN)

Managed data identifier ID: Depending on country or region, SPAIN_SOCIAL_SECURITY_NUMBER, USA_SOCIAL_SECURITY_NUMBER

Supported countries and regions: Spain, US

Keyword required: Yes. For Spain, keywords include: *número de la seguridad social, social security no., social security number, socialsecurityno#, ssn, ssn#*. For the US, keywords include: *social security, ss#, ssn*.

Comments: None

Taxpayer identification or reference number

Support includes: CUIL and CUIT codes for Argentina; CIF, NIE, and NIF numbers for Spain; CNPJ and CPF numbers for Brazil; Codice Fiscale numbers for Italy; ITINs for the US; NIT numbers for Colombia; PANs for India; RFC numbers for Mexico; RUN and RUT numbers for Chile; Steueridentifikationsnummer numbers for Germany; TFNs for Australia; TINs for France; and, TRN and UTR numbers for the UK.

Managed data identifier ID: Depending on country or region, ARGENTINA_INDIVIDUAL_TAX_IDENTIFICATION_NUMBER, ARGENTINA_ORGANIZATION_TAX_IDENTIFICATION_NUMBER, AUSTRALIA_TAX_FILE_NUMBER, BRAZIL_CNPJ_NUMBER, BRAZIL_CPF_NUMBER, CHILE_RUT_NUMBER, COLOMBIA_INDIVIDUAL_NIT_NUMBER, COLOMBIA_ORGANIZATION_NIT_NUMBER, FRANCE_TAX_IDENTIFICATION_NUMBER, GERMANY_TAX_IDENTIFICATION_NUMBER,

INDIA_PERMANENT_ACCOUNT_NUMBER, ITALY_NATIONAL_IDENTIFICATION_NUMBER, MEXICO_INDIVIDUAL_RFC_NUMBER, MEXICO_ORGANIZATION_RFC_NUMBER, SPAIN_NIE_NUMBER, SPAIN_NIF_NUMBER, SPAIN_TAX_IDENTIFICATION_NUMBER, UK_TAX_IDENTIFICATION_NUMBER, USA_INDIVIDUAL_TAX_IDENTIFICATION_NUMBER

Supported countries and regions: Argentina, Australia, Brazil, Chile, Colombia, France, Germany, India, Italy, Mexico, Spain, UK, US

Keyword required: Yes. The following table lists the keywords that Macie recognizes for specific countries and regions.

Country or region	Keywords
Argentina	argentina taxpayer id, clave única de identific ación tributaria, cuil, c.u.i.l, cuit, c.u.i.t, número de identificación fiscal, número de contribuy ente, unified labor identification code
Australia	tax file number, tfn
Brazil	cadastro de pessoa física, cadastro de pessoa fisica, cadastro de pessoas físicas, cadastro de pessoas fisicas, cadastro nacional da pessoa jurídica, cadastro nacional da pessoa juridica, cnpj, cpf
Chile	identidad número, nacional identidad, national unique role, nationaluniqueroleID#, número identificación, rol único nacional, rol único tributario, run, run#, r.u.n., rut, rut#, r.u.t., unique national number, unique national role, unique tax registry, unique tax role, unique tributary number, unique tributary role
Colombia	nit, nit., nit#, n.i.t.
France	numéro d'identification fiscal, tax id, tax identification number, tax number, tin, tin#

Country or region	Keywords
Germany	identifikationsnummer, steuer id, steueride ntifikationsnummer, steuernummer, tax id, tax identification number, tax number
India	e-pan, pan card, pan number, permanent account number
Italy	codice fiscal, dati anagrafici, ehic, health card, health insurance card, p. iva, partita i.v.a., personal data, tax code, tessera sanitaria
Mexico	código del registro federal de contribuyentes, identificación de impuestos, identificacion de impuestos, impuesto al valor agregado, iva, iva#, i.v.a., registro federal de contribuyentes, rfc, rfc#, r.f.c.
Spain	cif, cif número, cifnúmero#, nie, nif, número de contribuyente, número de identidad de extranjero, número de identificación fiscal, número de impuesto corporativo, personal tax number, tax id, tax identification number, tax number, tin, tin#
UK	paye, tax id, tax id no., tax id number, tax identification, tax identification#, tax no., tax number, tax reference, tax#, taxid#, temporary reference number, tin, trn, unique tax reference, unique taxpayer reference, utr
US	i.t.i.n., individual taxpayer identification number, itin

Comments: The managed data identifier for Chile (CHILE_RUT_NUMBER) is designed to detect both Rol Único Nacional (RUN) numbers and Rol Único Tributario (RUT) numbers. For Registro Federal de Contribuyentes (RFC) numbers for Mexico, Macie doesn't report occurrences

of the following character sequences, which are commonly used as example RFC numbers: `XAXX010101000` and `XEXX010101000`.

For several types of taxpayer identification and reference numbers, Macie doesn't report occurrences where all the digits are zeroes—for example, `00000000-K`, `000000000`, and `00.000.000`. This is because the use of only zeroes is common in examples of certain types of taxpayer identification and reference numbers.

Vehicle identification number (VIN)

Managed data identifier ID: VEHICLE_IDENTIFICATION_NUMBER

Supported countries and regions: Any, if the VIN is in proximity of a keyword in one of the following languages: English, French, German, Lithuanian, Polish, Portuguese, Romanian, or Spanish.

Keyword required: Yes. Keywords include: *Fahrgestellnummer, niv, numarul de identificare, numarul seriei de sasiu, numer VIN, Número de Identificação do Veículo, Número de Identificación de Automóviles, numéro d'identification du véhicule, vehicle identification number, vin, VIN numeris*

Comments: Macie can detect VINs that consist of a 17-character sequence and adhere to the ISO 3779 and 3780 standards. These standards were designed for worldwide use.

Building custom data identifiers

In addition to using the managed data identifiers that Amazon Macie provides, you can build and use custom data identifiers. A *custom data identifier* is a set of criteria that you define to detect sensitive data in Amazon Simple Storage Service (Amazon S3) objects. The criteria consist of a regular expression (*regex*) that defines a text pattern to match and, optionally, character sequences and a proximity rule that refine the results. The character sequences can be: *keywords*, which are words or phrases that must be in proximity of text that matches the regex, or *ignore words*, which are words or phrases to exclude from results.

With custom data identifiers, you can define detection criteria that reflect your organization's particular scenarios, intellectual property, or proprietary data. For example, you can detect employee IDs, customer account numbers, or internal data classifications. If you configure sensitive data discovery jobs or automated sensitive data discovery to use these identifiers, you can supplement the managed data identifiers that Macie provides.

In addition to detection criteria, you can optionally configure custom severity settings for findings that a custom data identifier produces. By default, Macie assigns the *Medium* severity to all the findings that a custom data identifier produces. Severity doesn't change based on the number of occurrences of text that match an identifier's detection criteria. If you configure custom severity settings, severity can be based on the number of occurrences of text that match the criteria.

Topics

Configuration options for custom data identifiers

By using custom data identifiers, you can define custom criteria for detecting sensitive data in Amazon Simple Storage Service (Amazon S3) objects. You can supplement the managed data identifiers that Amazon Macie provides, and detect sensitive data that reflects your organization's particular scenarios, intellectual property, or proprietary data.

Each custom data identifier specifies detection criteria and, optionally, severity settings for findings that the identifier produces. The detection criteria specify a regular expression that defines a text pattern to match in an S3 object. The criteria can also specify character sequences and a proximity rule that refine the results. The severity settings specify which severity to assign to findings. Severity can be based on the number of occurrences of text that match the identifier's detection criteria.

Topics

Detection criteria

When you create a custom data identifier, you specify a regular expression (*regex*) that defines a text pattern to match. You can also specify character sequences, such as words and phrases, and a proximity rule that refine the results. The character sequences can be: *keywords*, which are words or phrases that must be in proximity of text that matches the regex, or *ignore words*, which are words or phrases to exclude from results.

For the regex, Amazon Macie supports a subset of the pattern syntax provided by the Perl Compatible Regular Expressions (PCRE) library. Of the constructs provided by the PCRE library, Macie doesn't support the following pattern elements:

- Backreferences
- Capturing groups
- Conditional patterns
- Embedded code
- Global pattern flags, such as `/i`, `/m`, and `/x`
- Recursive patterns
- Positive and negative look-behind and look-ahead zero-width assertions, such as `?=`, `?!`, `?<=`, and `?<!`

The regex can contain as many as 512 characters.

To create an effective regex pattern for a custom data identifier, note the following tips and recommendations:

- Use anchors (^ or $) only if you expect the pattern to appear at the beginning or end of a file, not the beginning or end of a line.
- For performance reasons, Macie limits the size of bounded repeat groups. For example, `\d{100,1000}` won't compile in Macie. To approximate this functionality, you can use an open-ended repeat such as `\d{100,}`.
- To make parts of a pattern case insensitive, you can use the `(?i)` construct instead of the `/i` flag.
- There's no need to optimize prefixes or alternations manually. For example, changing `/hello|hi|hey/` to `/h(?:ello|i|ey)/` won't improve performance.
- For performance reasons, Macie limits the number of repeated wildcards. For example, `a*b*a*` won't compile in Macie.

To protect against malformed or long-running expressions, Macie automatically tests regex patterns against a collection of sample text when you create a custom data identifier. If there's an issue with the regex, Macie returns an error that describes the issue.

In addition to the regex, you can optionally specify character sequences and a proximity rule to refine the results.

Keywords

These are specific character sequences that must be in proximity of text that matches the regex pattern. The proximity requirements vary based on an S3 object's storage format or file type:

- **Structured columnar data** – Macie includes a result if the text matches the regex pattern and a keyword is in the name of the field or column that stores the text, or the text is preceded by and within the maximum match distance of a keyword in the same field or cell value. This is the case for Microsoft Excel workbooks, CSV files, and TSV files.
- **Structured record-based data** – Macie includes a result if the text matches the regex pattern and the text is within the maximum match distance of a keyword. The keyword can be in the name of an element in the path to the field or array that stores the text, or it can precede and be part of the same value in the field or array that stores the text. This is the case for Apache Avro object containers, Apache Parquet files, JSON files, and JSON Lines files.
- **Unstructured data** – Macie includes a result if the text matches the regex pattern and the text is preceded by and within the maximum match distance of a keyword. This is the case for Adobe Portable Document Format files, Microsoft Word documents, email messages, and non-binary text files other than CSV, JSON, JSON Lines, and TSV files. This includes any structured data, such as tables, in these types of files.

You can specify as many as 50 keywords. Each keyword can contain 3–90 UTF-8 characters. Keywords aren't case sensitive.

Maximum match distance

This is a character-based proximity rule for keywords. Macie uses this setting to determine whether a keyword precedes text that matches the regex pattern. The setting defines the maximum number of characters that can exist between the end of a complete keyword and the end of text that matches the regex pattern. Macie includes a result if the text:

- Matches the regex pattern,
- Occurs after at least one complete keyword, and
- Occurs within the specified distance of the keyword.

Otherwise, Macie excludes the text from results.

You can specify a distance of 1–300 characters. The default distance is 50 characters. For best results, this distance should be greater than the minimum number of characters of text that the

regex is designed to detect. If only part of the text is within the maximum match distance of a keyword, Macie doesn't include it in results.

Ignore words

These are specific character sequences to exclude from results. If text matches the regex pattern but it contains an ignore word, Macie doesn't include it in results.

You can specify as many as 10 ignore words. Each ignore word can contain 4–90 UTF-8 characters. Ignore words are case sensitive.

> **Note**
>
> Before you create a custom data identifier, we strongly recommend that you test and refine its detection criteria with sample data. Because custom data identifiers are used by sensitive data discovery jobs, you can't change a custom data identifier after you create it. This helps ensure that you have an immutable history of sensitive data findings and discovery results for data privacy and protection audits or investigations that you perform. You can test detection criteria by using the Amazon Macie console or the Amazon Macie API. To test the criteria by using the console, use the options in the **Evaluate** section while you're creating the custom data identifier. To test the criteria programmatically, use the TestCustomDataIdentifier operation of the Amazon Macie API. If you're using the AWS Command Line Interface, run the test-custom-data-identifier command to test the criteria.

For a demonstration of how keywords can help you find sensitive data and avoid false positives, watch the following video: How Amazon Macie uses keywords to discover sensitive data.

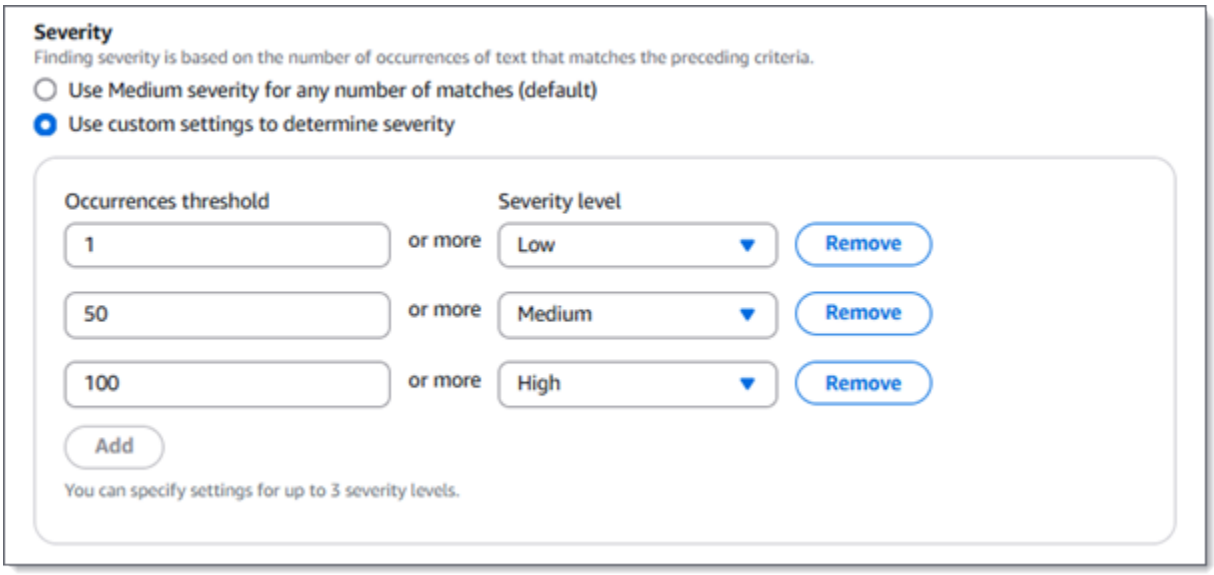
Severity settings for findings

When you create a custom data identifier, you can also specify custom severity settings for sensitive data findings that the identifier produces. By default, Amazon Macie assigns the *Medium* severity to all the findings that a custom data identifier produces. If an S3 object contains at least one occurrence of text that matches the detection criteria, Macie automatically assigns the *Medium* severity to the resulting finding.

With custom severity settings, you specify which severity to assign based on the number of occurrences of text that match the detection criteria. You can define *occurrences thresholds* for as

many as three severity levels: *Low* (least severe), *Medium*, and *High* (most severe). An *occurrences threshold* is the minimum number of matches that must exist in an S3 object to produce a finding with the specified severity. If you specify more than one threshold, the thresholds must be in ascending order by severity, moving from *Low* to *High*.

For example, the following image shows severity settings that specify three occurrences thresholds, one for each severity level that Macie supports.



The following table indicates the severity of the findings that the custom data identifier produces.

Occurrences threshold	Severity level	Result
1	Low	If an S3 object contains 1–49 occurrences of text that match the detection criteria, the severity of the resulting finding is *Low*.
50	Medium	If an S3 object contains 50–99 occurrences of text that match the detection criteria, the severity of the resulting finding is *Medium*.
100	High	If an S3 object contains 100 or more occurrences of text

Occurrences threshold	Severity level	Result
		that match the detection criteria, the severity of the resulting finding is *High*.

You can also use severity settings to specify whether to create a finding at all. If an S3 object contains fewer occurrences than the lowest occurrences threshold, Macie doesn't create a finding.

Creating a custom data identifier

A *custom data identifier* is a set of criteria that you define to detect sensitive data in Amazon Simple Storage Service (Amazon S3) objects. When you create a custom data identifier, you specify a regular expression (*regex*) that defines a text pattern to match in an S3 object. You can also specify character sequences and a proximity rule that refine the results. The character sequences can be: *keywords*, which are words or phrases that must be in proximity of text that matches the regex, or *ignore words*, which are words or phrases to exclude from results. By using custom data identifiers, you can supplement the managed data identifiers that Amazon Macie provides, and detect sensitive data that reflects your organization's particular scenarios, intellectual property, or proprietary data.

For example, many companies have a specific syntax for employee IDs. One such syntax might be: a capital letter that indicates whether an employee is a full-time (*F*) or part-time (*P*) employee, followed by a hyphen (–), followed by an eight-digit sequence that identifies the employee. Examples are: *F–12345678* for a full-time employee, and *P–87654321* for a part-time employee. To detect employee IDs that use this syntax, you might create a custom data identifier that specifies the following regex: `[A-Z]-\d{8}`. To refine the analysis and avoid false positives, you might also configure the identifier to use keywords (`employee` and `employee ID`) and a maximum match distance of 20 characters. With these criteria, results include text that matches the regex if the text occurs after the keyword *employee* or *employee ID* and all the text occurs within 20 characters of one of those keywords.

For a demonstration of how keywords can help you find sensitive data and avoid false positives, watch the following video: How Amazon Macie uses keywords to discover sensitive data.

In addition to detection criteria, you can optionally specify custom severity settings for findings that a custom data identifier produces. Severity can be based on the number of occurrences

of text that match the identifier's detection criteria. If you don't specify these settings, Macie automatically assigns the *Medium* severity to all the findings that the identifier produces. Severity doesn't change based on the number of occurrences of text that match the identifier's detection criteria.

For detailed information about these and other settings, see Configuration options for custom data identifiers.

To create a custom data identifier

You can create a custom data identifier by using the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to create a custom data identifier by using the Amazon Macie console.

To create a custom data identifier

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, under **Settings**, choose **Custom data identifiers**.
3. Choose **Create**.
4. For **Name**, enter a name for the custom data identifier. The name can contain as many as 128 characters.
5. For **Description**, optionally enter a brief description of the custom data identifier. The description can contain as many as 512 characters.

   > **Note**
   >
   > Avoid including sensitive data in the name or description of a custom data identifier. Other users of your account might be able to access the name or description, depending on the actions that they're allowed to perform in Macie.

6. For **Regular expression**, enter the regular expression (*regex*) that defines the text pattern to match. The regex can contain as many as 512 characters.

   Macie supports a subset of the pattern syntax provided by the Perl Compatible Regular Expressions (PCRE) library. For additional details and tips, see Detection criteria for custom data identifiers.

7. For **Keywords**, optionally enter as many as 50 character sequences (separated by commas) to define specific text that must be in proximity of text that matches the regex pattern.

   Macie includes an occurrence in results only if the text matches the regex pattern and the text is within the maximum match distance of one of these keywords. Each keyword can contain 3–90 UTF-8 characters. Keywords aren't case sensitive.

8. For **Ignore words**, optionally enter as many as 10 character sequences (separated by commas) that define specific text to exclude from results.

   Macie excludes an occurrence from results if the text matches the regex pattern but it contains one of these ignore words. Each ignore word can contain 4–90 UTF-8 characters. Ignore words are case sensitive.

9. For **Maximum match distance**, optionally enter the maximum number of characters that can exist between the end of a keyword and the end of text that matches the regex pattern.

   Macie includes an occurrence in results only if the text matches the regex pattern and the text is within this distance of a complete keyword. The distance can be 1–300 characters. The default distance is 50 characters.

10. For **Severity**, choose how to determine the severity of sensitive data findings that the custom data identifier produces:

    - To automatically assign the *Medium* severity to all findings, choose **Use Medium severity for any number of matches (default)**. With this option, Macie automatically assigns the *Medium* severity to a finding if the affected S3 object contains one or more occurrences of text that match the detection criteria.
    - To assign severity based on occurrences thresholds that you specify, choose **Use custom settings to determine severity**. Then use the **Occurrences threshold** and **Severity level** options to specify the minimum number of matches that must exist in an S3 object to produce a finding with a selected severity.

      You can specify as many as three occurrences thresholds, one for each severity level that Macie supports: *Low* (least severe), *Medium*, or *High* (most severe). If you specify more than one, the thresholds must be in ascending order by severity, moving from *Low* to *High*. If an S3 object contains fewer occurrences than the lowest threshold, Macie doesn't create a finding.

11. (Optional) For **Tags**, choose **Add tag**, and then enter as many as 50 tags to assign to the custom data identifier.

    A *tag* is a label that you define and assign to certain types of AWS resources. Each tag consists of a required tag key and an optional tag value. Tags can help you identify, categorize, and manage resources in different ways, such as by purpose, owner, environment, or other criteria. To learn more, see Tagging Macie resources.

12. (Optional) For **Evaluate**, enter up to 1,000 characters in the **Sample data** box, and then choose **Test** to test the detection criteria. Macie evaluates the sample data and reports the number of occurrences of text that match the criteria. You can repeat this step as many times as you like to refine and optimize the criteria.

    > **Note**
    >
    > We strongly recommend that you test and refine the detection criteria with sample data. Because custom data identifiers are used by sensitive data discovery jobs, you can't change a custom data identifier after you create it. This helps ensure that you have an immutable history of sensitive data findings and discovery results.

13. When you finish, choose **Submit**.

Macie tests the settings and verifies that it can compile the regex. If there's an issue with a setting or the regex, Macie displays an error that describes the issue. After you address any issues, you can save the custom data identifier.

API

To create a custom data identifier programmatically, use the CreateCustomDataIdentifier operation of the Amazon Macie API. Or, if you're using the AWS Command Line Interface (AWS CLI), run the create-custom-data-identifier command.

> **Note**
>
> Before you create a custom data identifier, we strongly recommend that you test and refine its detection criteria with sample data. Because custom data identifiers are used by sensitive data discovery jobs, you can't change a custom data identifier after you create it. This helps ensure that you have an immutable history of sensitive data findings and discovery results.

To test the criteria programmatically, you can use the TestCustomDataIdentifier operation of the Amazon Macie API. This operation provides an environment for evaluating sample data with detection criteria. If you're using the AWS CLI, you can run the test-custom-data-identifier command to test the criteria.

When you're ready to create the custom data identifier, use the following parameters to define its detection criteria:

- `regex` – Specify the regular expression (*regex*) that defines the text pattern to match. The regex can contain as many as 512 characters.

  Macie supports a subset of the pattern syntax provided by the Perl Compatible Regular Expressions (PCRE) library. For additional details and tips, see Detection criteria for custom data identifiers.

- `keywords` – Optionally specify 1–50 character sequences (*keywords*) that must be in proximity of text that matches the regex pattern.

  Macie includes an occurrence in results only if the text matches the regex pattern and the text is within the maximum match distance of one of these keywords. Each keyword can contain 3–90 UTF-8 characters. Keywords aren't case sensitive.

- `maximumMatchDistance` – Optionally specify the maximum number of characters that can exist between the end of a keyword and the end of text that matches the regex pattern. If you're using the AWS CLI, use the `maximum-match-distance` parameter to specify this value.

  Macie includes an occurrence in results only if the text matches the regex pattern and the text is within this distance of a complete keyword. The distance can be 1–300 characters. The default distance is 50 characters.

- `ignoreWords` – Optionally specify 1–10 character sequences (*ignore words*) to exclude from results. If you're using the AWS CLI, use the `ignore-words` parameter to specify these character sequences.

  Macie excludes an occurrence from results if the text matches the regex pattern but it contains one of these ignore words. Each ignore word can contain 4–90 UTF-8 characters. Ignore words are case sensitive.

To specify the severity of sensitive data findings that the custom data identifier produces, use the `severityLevels` parameter or, if you're using the AWS CLI, the `severity-levels` parameter:

- To automatically assign the `MEDIUM` severity to all the findings, omit this parameter. Macie then uses the default setting. By default, Macie assigns the `MEDIUM` severity to a finding if the affected S3 object contains one or more occurrences of text that match the detection criteria.
- To assign severity based on occurrences thresholds that you specify, specify the minimum number of matches that must exist in an S3 object to produce a finding with a specified severity.

  You can specify as many as three occurrences thresholds, one for each severity level that Macie supports: `LOW` (least severe), `MEDIUM`, or `HIGH` (most severe). If you specify more than one, the thresholds must be in ascending order by severity, moving from `LOW` to `HIGH`. If an S3 object contains fewer occurrences than the lowest threshold, Macie doesn't create a finding.

Use additional parameters to specify a name and other settings, such as tags, for the custom data identifier. Avoid including sensitive data in these settings. Other users of your account might be able to access these values, depending on the actions that they're allowed to perform in Macie.

When you submit your request, Macie tests the settings and verifies that it can compile the regex. If there's an issue with a setting or the regex, the request fails and Macie returns a message that describes the issue. If the request succeeds, you receive output similar to the following:

```
{
    "customDataIdentifierId": "393950aa-82ea-4bdc-8f7b-e5be3example"
}
```

Where `customDataIdentifierId` specifies the unique identifier (ID) for the custom data identifier that was created.

To subsequently retrieve and review the settings for the custom data identifier, use the GetCustomDataIdentifier operation or, if you're using the AWS CLI, run the get-custom-data-identifier command. For the `id` parameter, specify the custom data identifier's ID.

The following examples show how to use the AWS CLI to create a custom data identifier. The examples create a custom data identifier that's designed to detect employee IDs that use a specific syntax and are within proximity of a specified keyword. The examples also define custom severity settings for findings that the identifier produces.

This example is formatted for Linux, macOS, or Unix, and it uses the backslash (\) line-continuation character to improve readability.

```
$ aws macie2 create-custom-data-identifier \
--name "EmployeeIDs" \
--regex "[A-Z]-\d{8}" \
--keywords '["employee","employee ID"]' \
--maximum-match-distance 20 \
--severity-levels '[{"occurrencesThreshold":1,"severity":"LOW"},
{"occurrencesThreshold":50,"severity":"MEDIUM"},
{"occurrencesThreshold":100,"severity":"HIGH"}]' \
--description "Detects employee IDs in proximity of a keyword." \
--tags '{"Stack":"Production"}'
```

This example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws macie2 create-custom-data-identifier ^
--name "EmployeeIDs" ^
--regex "[A-Z]-\d{8}" ^
--keywords "[\"employee\",\"employee ID\"]" ^
--maximum-match-distance 20 ^
--severity-levels "[{\"occurrencesThreshold\":1,\"severity\":\"LOW\"},
{\"occurrencesThreshold\":50,\"severity\":\"MEDIUM\"},{\"occurrencesThreshold\":100,
\"severity\":\"HIGH\"}]" ^
--description "Detects employee IDs in proximity of a keyword." ^
--tags={\"Stack\":\"Production\"}
```

Where:

- *EmployeeIDs* is the name of the custom data identifier.
- *[A-Z]-\d{8}* is the regex for the text pattern to match.
- *employee* and *employee ID* are keywords that must be in proximity of text that matches the regex pattern.

- *20* is the maximum number of characters that can exist between the end of a keyword and the end of text that matches the regex pattern.
- `description` specifies a brief description of the custom data identifier.
- `severity-levels` defines custom occurrences thresholds for the severity of findings that the custom data identifier produces: *LOW* for 1–49 occurrences; *MEDIUM* for 50–99 occurrences; and, *HIGH* for 100 or more occurrences.
- *Stack* is the tag key of the tag to assign to the custom data identifier. *Production* is the tag value for the specified tag key.

After you create the custom data identifier, you can create and configure sensitive data discovery jobs to use it, or add it to your settings for automated sensitive data discovery.

Deleting a custom data identifier

After you create a custom data identifier, you can delete it. If you do this, Amazon Macie soft deletes the custom data identifier. This means that a record of the custom data identifier remains for your account, but it's marked as deleted. If a custom data identifier has this status, you can't configure new sensitive data discovery jobs to use it or add it to your settings for automated sensitive data discovery. In addition, you can no longer access it by using the Amazon Macie console. You can, however, retrieve its settings by using the Amazon Macie API. If you delete a custom data identifier, it doesn't count against the quota of custom data identifiers for your account.

If you configure a sensitive data discovery job to use a custom data identifier that you subsequently delete, the job will run as scheduled and continue to use the custom data identifier. This means that your job results, both sensitive data findings and sensitive data discovery results, will report text that matches the identifier's criteria. This helps ensure that you have an immutable history of sensitive data findings and discovery results for data privacy and protection audits or investigations that you perform.

Similarly, if you configure automated sensitive data discovery to use a custom data identifier that you subsequently delete, daily analysis cycles will proceed and continue to use the custom data identifier. This means that sensitive data findings, statistics, and other types of results will continue to report text that matches the identifier's criteria.

Before you delete a custom data identifier, do the following to prevent Macie from using it during subsequent analysis cycles and job runs:

- Check your settings for automated sensitive data discovery. If you added the custom data identifier to these settings, remove it. For more information, see Configuring settings for automated sensitive data discovery.
- Review your job inventory to identify jobs that use the custom data identifier and are scheduled to run in the future. If you want a job to stop using the custom data identifier, you can cancel the job. Then create a copy of the job, adjust the settings for the copy, and save the copy as a new job. For more information, see Managing sensitive data discovery jobs.

It's also a good idea to note the unique identifier (ID) that Macie assigned to the custom data identifier. You'll need this ID if you later want to review the custom data identifier's settings.

After you complete the preceding tasks, delete the custom data identifier.

To delete a custom data identifier

You can delete a custom data identifier by using the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to delete a custom data identifier by using the Amazon Macie console.

To delete a custom data identifier

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, under **Settings**, choose **Custom data identifiers**.
3. To note the unique identifier (ID) for the custom data identifier that you want to delete, choose the custom data identifier's name. On the page that appears, the **Id** box displays this ID. After you note the ID, choose **Custom data identifiers** in the navigation pane again.
4. On the **Custom data identifiers** page, select the checkbox for the custom data identifier to delete.
5. On the **Actions** menu, choose **Delete**.
6. When prompted for confirmation, choose **Ok**.

API

To delete a custom data identifier programmatically, use the DeleteCustomDataIdentifier operation of the Amazon Macie API. Or, if you're using the AWS Command Line Interface (AWS CLI), run the delete-custom-data-identifier command.

For the `id` parameter, specify the unique identifier (ID) for the custom data identifier that you want to delete. You can get this ID by using the ListCustomDataIdentifiers operation. This operation retrieves a subset of information about the custom data identifiers for your account. If you're using the AWS CLI, you can run the list-custom-data-identifiers command to retrieve this information.

The following example shows how to delete a custom data identifier by using the AWS CLI.

```
$ aws macie2 delete-custom-data-identifier --id 393950aa-82ea-4bdc-8f7b-e5be3example
```

Where *393950aa-82ea-4bdc-8f7b-e5be3example* is the ID for the custom data identifier to delete.

If the request succeeds, Macie returns an empty HTTP 200 response. Otherwise, Macie returns an HTTP 4*xx* or 500 response indicating why the request failed.

To review a custom data identifier's settings after you delete it, use the GetCustomDataIdentifier operation of the Amazon Macie API. Or, if you're using the AWS CLI, run the get-custom-data-identifier command. For the `id` parameter, specify the custom data identifier's ID. After you delete a custom data identifier, you can't access its settings by using the Amazon Macie console.

Defining sensitive data exceptions with allow lists

With allow lists in Amazon Macie, you can define specific text and text patterns that you want Macie to ignore when it inspects Amazon Simple Storage Service (Amazon S3) objects for sensitive data. These are typically sensitive data exceptions for your particular scenarios or environment. If data matches text or a text pattern in an allow list, Macie doesn't report the data. This is the case even if the data matches the criteria of a managed data identifier or a custom data identifier. By using allow lists, you can refine your analysis of Amazon S3 data and reduce noise.

You can create and use two types of allow lists in Macie:

- **Predefined text** – For this type of list, you specify certain character sequences to ignore. For example, you might specify the names of public representatives for your organization, specific phone numbers, or specific sample data that your organization uses for testing. If you use this type of list, Macie ignores text that exactly matches an entry in the list.

  This type of allow list is helpful if you want to specify words, phrases, and other kinds of character sequences that aren't sensitive, aren't likely to change, and don't necessarily adhere to a common pattern.
- **Regular expression** – For this type of list, you specify a regular expression (*regex*) that defines a text pattern to ignore. For example, you might specify the pattern for your organization's public phone numbers, email addresses for your organization's domain, or patterned sample data that your organization uses for testing. If you use this type of list, Macie ignores text that completely matches the pattern defined by the list.

  This type of allow list is helpful if you want to specify text that isn't sensitive but varies or is likely to change while also adhering to a common pattern.

After you create an allow list, you can create and configure sensitive data discovery jobs to use it, or add it to your settings for automated sensitive data discovery. Macie then uses the list when it analyzes data. If Macie finds text that matches an entry or pattern in an allow list, Macie doesn't report that occurrence of text in sensitive data findings, statistics, and other types of results.

You can manage and use allow lists in all the AWS Regions where Macie is currently available except the Asia Pacific (Osaka) Region.

Topics

Configuration options and requirements for allow lists

In Amazon Macie, you can use allow lists to specify text or text patterns that you want Macie to ignore when it inspects Amazon Simple Storage Service (Amazon S3) objects for sensitive data. Macie provides options for two types of allow lists, predefined text and regular expressions.

A list of predefined text is helpful if you want Macie to ignore specific words, phrases, and other kinds of character sequences that you don't consider sensitive. Examples are: the names of public representatives for your organization, specific phone numbers, or specific sample data that your organization uses for testing. If Macie finds text that matches the criteria of a managed or custom data identifier and the text also matches an entry in an allow list, Macie doesn't report that occurrence of text in sensitive data findings, statistics, and other types of results.

A regular expression (*regex*) is helpful if you want Macie to ignore text that varies or is likely to change while also adhering to a common pattern. The regex specifies a text pattern to ignore. Examples are: public phone numbers for your organization, email addresses for your organization's domain, or patterned sample data that your organization uses for testing. If Macie finds text that matches the criteria of a managed or custom data identifier and the text also matches a regex pattern in an allow list, Macie doesn't report that occurrence of text in sensitive data findings, statistics, and other types of results.

You can create and use both types of allow lists in all the AWS Regions where Macie is currently available except the Asia Pacific (Osaka) Region. As you create and manage allow lists, keep the following options and requirements in mind. Also note that list entries and regex patterns for mailing addresses aren't supported.

Topics

Options and requirements for lists of predefined text

For this type of allow list, you provide a line-delimited plaintext file that lists specific character sequences to ignore. The list entries are typically words, phrases, and other kinds of character sequences that you don't consider sensitive, aren't likely to change, and don't necessarily adhere to a specific pattern. If you use this type of list, Amazon Macie doesn't report occurrences of text that exactly match an entry in the list. Macie treats each list entry as a string literal value.

To use this type of allow list, start by creating the list in a text editor and saving it as a plaintext file. Then upload the list to an S3 general purpose bucket. Also ensure that the storage and encryption settings for the bucket and the object allow Macie to retrieve and decrypt the list. Then create and configure settings for the list in Macie.

After you configure the settings in Macie, we recommend that you test the allow list with a small, representative set of data for your account or organization. To test a list, you can create a one-time job. Configure the job to use the list in addition to the managed and custom data identifiers that you typically use to analyze data. You can then review the job's results—sensitive data findings, sensitive data discovery results, or both. If the job's results differ from what you expect, you can change and test the list until the results are what you expect.

After you finish configuring and testing an allow list, you can create and configure additional jobs to use it, or add it to your settings for automated sensitive data discovery. When those jobs start to run or the next automated discovery analysis cycle starts, Macie retrieves the latest version of the list from Amazon S3 and stores it in temporary memory. Macie then uses this temporary copy of the list when it inspects S3 objects for sensitive data. When a job finishes running or the analysis cycle is complete, Macie permanently deletes its copy of the list from memory. The list doesn't persist in Macie. Only the list's settings persist in Macie.

> **Important**
>
> Because lists of predefined text don't persist in Macie, it's important to check the status of your allow lists periodically. If Macie can't retrieve or parse a list that you configured a job or automated discovery to use, Macie doesn't use the list. This might produce unexpected results, such as sensitive data findings for text that you specified in the list.

Topics

- Syntax requirements

- Storage requirements
- Encryption/Decryption requirements
- Design considerations and recommendations

Syntax requirements

When you create this type of allow list, note the following requirements for the list's file:

- The list must be stored as a plaintext (`text/plain`) file, such as a .txt, .text, or .plain file.
- The list must use line breaks to separate individual entries. For example:

```
Akua Mansa
John Doe
Martha Rivera
425-555-0100
425-555-0101
425-555-0102
```

  Macie treats each line as a single, distinct entry in the list. The file can also contain blank lines to improve readability. Macie skips blank lines when it parses the file.
- Each entry can contain 1–90 UTF–8 characters.
- Each entry must be a complete, exact match for the text to ignore. Macie doesn't support use of wildcard characters or partial values for entries. Macie treats each entry as a string literal value. Matches aren't case sensitive.
- The file can contain 1–100,000 entries.
- The total storage size of the file can't exceed 35 MB.

Storage requirements

As you add and manage allow lists in Amazon S3, note the following storage requirements and recommendations:

- **Regional support** – An allow list must be stored in a bucket that's in the same AWS Region as your Macie account. Macie can’t access an allow list if it’s stored in a different Region.
- **Bucket ownership** – An allow list must be stored in a bucket that's owned by your AWS account. If you want other accounts to use the same allow list, consider creating an Amazon S3 replication

rule to replicate the list to buckets that are owned by those accounts. For information about replicating S3 objects, see Replicating objects in the *Amazon Simple Storage Service User Guide*.

In addition, your AWS Identity and Access Management (IAM) identity must have read access to the bucket and object that store the list. Otherwise, you won't be allowed to create or update the list's settings or check the list's status by using Macie.

- **Storage types and classes** – An allow list must be stored in a general purpose bucket, not a directory bucket. In addition, it must be stored using one of the following storage classes: Reduced Redundancy (RRS), S3 Glacier Instant Retrieval, S3 Intelligent-Tiering, S3 One Zone-IA, S3 Standard, or S3 Standard-IA.
- **Bucket policies** – If you store an allow list in a bucket that has a restrictive bucket policy, ensure that the policy allows Macie to retrieve the list. To do this, you can add a condition for the Macie service-linked role to the bucket policy. For more information, see Allowing Macie to access S3 buckets and objects.

  Also ensure that the policy allows your IAM identity to have read access to the bucket. Otherwise, you won't be allowed to create or update the list's settings or check the list's status by using Macie.
- **Object paths** – If you store more than one allow list in Amazon S3, the object path for each list must be unique. In other words, each allow list must be stored separately in its own S3 object.
- **Versioning** – When you add an allow list to a bucket, we recommend that you also enable versioning for the bucket. You can then use date and time values to correlate versions of the list with the results of sensitive data discovery jobs and automated sensitive data discovery cycles that use the list. This can help with data privacy and protection audits or investigations that you perform.
- **Object Lock** – To prevent an allow list from being deleted or overwritten for a certain amount of time or indefinitely, you can enable Object Lock for the bucket that stores the list. Enabling this setting doesn't prevent Macie from accessing the list. For information about this setting, see Locking objects with Object Lock in the *Amazon Simple Storage Service User Guide*.

Encryption/Decryption requirements

If you encrypt an allow list in Amazon S3, the permissions policy for the Macie service-linked role typically grants Macie the permissions that it needs to decrypt the list. However, this depends on the type of encryption that's used:

- If a list is encrypted using server-side encryption with an Amazon S3 managed key (SSE-S3), Macie can decrypt the list. The service-linked role for your Macie account grants Macie the permissions that it needs.
- If a list is encrypted using server-side encryption with an AWS managed AWS KMS key (DSSE-KMS or SSE-KMS), Macie can decrypt the list. The service-linked role for your Macie account grants Macie the permissions that it needs.
- If a list is encrypted using server-side encryption with a customer managed AWS KMS key (DSSE-KMS or SSE-KMS), Macie can decrypt the list only if you allow Macie to use the key. To learn how to do this, see Allowing Macie to use a customer managed AWS KMS key.

  > **Note**
  >
  > You can encrypt a list with a customer managed AWS KMS key in an external key store. However, the key might then be slower and less reliable than a key that's managed entirely within AWS KMS. If latency or an availability issue prevents Macie from decrypting the list, Macie doesn't use the list when it analyzes S3 objects. This might produce unexpected results, such as sensitive data findings for text that you specified in the list. To reduce this risk, consider storing the list in an S3 bucket that's configured to use the key as an S3 Bucket Key.
  >
  > For information about using KMS keys in external key stores, see External key stores in the *AWS Key Management Service Developer Guide*. For information about using S3 Bucket Keys, see Reducing the cost of SSE-KMS with Amazon S3 Bucket Keys in the *Amazon Simple Storage Service User Guide*.

- If a list is encrypted using server-side encryption with a customer-provided key (SSE-C) or client-side encryption, Macie can't decrypt the list. Consider using SSE-S3, DSSE-KMS, or SSE-KMS encryption instead.

If a list is encrypted with an AWS managed KMS key or a customer managed KMS key, your AWS Identity and Access Management (IAM) identity must also be allowed to use the key. Otherwise, you won't be allowed to create or update the list's settings or check the list's status by using Macie. To learn how to check or change the permissions for a KMS key, see Key policies in AWS KMS in the *AWS Key Management Service Developer Guide*.

For detailed information about encryption options for Amazon S3 data, see Protecting data with encryption in the *Amazon Simple Storage Service User Guide*.

Design considerations and recommendations

In general, Macie treats each entry in an allow list as a string literal value. That is to say, Macie ignores each occurrence of text that exactly matches a complete entry in an allow list. Matches aren't case sensitive.

However, Macie uses the entries as part of a larger data extraction and analysis framework. The framework includes machine learning and pattern matching functions that factor dimensions such as grammatical and syntactical variations and, in many cases, keyword proximity. The framework also factors an S3 object's file type or storage format. Therefore, keep the following considerations and recommendations in mind as you add and manage the entries in an allow list.

Prepare for different file types and storage formats

For unstructured data, such as text in an Adobe Portable Document Format (.pdf) file, Macie ignores text that exactly matches a complete entry in an allow list, including text that spans multiple lines or pages.

For structured data, such as columnar data in a CSV file or record-based data in a JSON file, Macie ignores text that exactly matches a complete entry in an allow list if all the text is stored in a single field, cell, or array. This requirement doesn't apply to structured data that's stored in an otherwise unstructured file, such as a table in a .pdf file.

For example, consider the following content in a CSV file:

```
Name,Account ID
Akua Mansa,111111111111
John Doe,222222222222
```

If `Akua Mansa` and `John Doe` are entries in an allow list, Macie ignores those names in the CSV file. The complete text of each list entry is stored in a single `Name` field.

Conversely, consider a CSV file that contains the following columns and fields:

```
First Name,Last Name,Account ID
Akua,Mansa,111111111111
John,Doe,222222222222
```

If `Akua Mansa` and `John Doe` are entries in an allow list, Macie doesn't ignore those names in the CSV file. None of the fields in the CSV file contain the complete text of an entry in the allow list.

Include common variations

Add entries for common variations of numeric data, proper nouns, terms, and alphanumeric character sequences. For example, if you add names or phrases that contain only one space between words, also add variations that include two spaces between words. Similarly, add words and phrases that do and don't contain special characters, and consider including common syntactical and semantic variations.

For the US phone number *425-555-0100*, for example, you might add these entries to an allow list:

```
425-555-0100
425.555.0100
(425) 555-0100
+1-425-555-0100
```

For the date *February 1, 2022* in a multinational context, you might add entries that include common syntactical variations for English and French, including variations that do and don't include special characters:

```
February 1, 2022
1 février 2022
1 fevrier 2022
Feb 01, 2022
1 fév 2022
1 fev 2022
02/01/2022
01/02/2022
```

For names of people, include entries for various forms of a name that you don't consider sensitive. For example, include: the first name followed by the last name; the last name followed by the first name, the first and last name separated by one space; the first and last name separated by two spaces; and nicknames.

For the name *Martha Rivera*, for example, you might add:

```
Martha Rivera
Martha  Rivera
Rivera, Martha
Rivera,  Martha
Rivera Martha
```

```
Rivera  Martha
```

If you want to ignore variations of a specific name that contains many parts, create an allow list that uses a regular expression instead. For example, for the name *Dr. Martha Lyda Rivera, PhD*, you might use the following regular expression: `^(Dr. )?Martha\s(Lyda|L\.)?\s?Rivera,?( PhD)?$`.

Options and requirements for regular expressions

For this type of allow list, you specify a regular expression (*regex*) that defines a text pattern to ignore. For example, you might specify the pattern for your organization's public phone numbers, email addresses for your organization's domain, or patterned sample data that your organization uses for testing. The regex defines a common pattern for a specific kind of data that you don't consider sensitive. If you use this type of allow list, Amazon Macie doesn't report occurrences of text that completely match the specified pattern. Unlike an allow list that specifies predefined text to ignore, you create and store the regex and all other list settings in Macie.

When you create or update this type of allow list, you can test the list's regex with sample data before you save the list. We recommend that you do this with multiple sets of sample data. If you create a regex that's too general, Macie might ignore occurrences of text that you consider sensitive. If a regex is too specific, Macie might not ignore occurrences of text that you don't consider sensitive. To protect against malformed or long-running expressions, Macie also compiles and tests the regex against a collection of sample text automatically, and notifies you of issues to address.

For additional testing, we recommend that you also test the list's regex with a small, representative set of data for your account or organization. To do this, you can create a one-time job. Configure the job to use the list in addition to the managed and custom data identifiers that you typically use to analyze data. You can then review the job's results—sensitive data findings, sensitive data discovery results, or both. If the job's results differ from what you expect, you can change and test the regex until the results are what you expect.

After you configure and test an allow list, you can create and configure additional jobs to use it, or add it to your settings for automated sensitive data discovery. When those job run or Macie performs automated discovery, Macie uses the latest version of the list's regex to analyze data.

Topics

- Syntax support and recommendations

- Examples

Syntax support and recommendations

An allow list can specify a regular expression (*regex*) that contains as many as 512 characters. Macie supports a subset of the regex pattern syntax provided by the Perl Compatible Regular Expressions (PCRE) library. Of the constructs provided by the PCRE library, Macie doesn't support the following pattern elements:

- Backreferences
- Capturing groups
- Conditional patterns
- Embedded code
- Global pattern flags, such as `/i`, `/m`, and `/x`
- Recursive patterns
- Positive and negative look-behind and look-ahead zero-width assertions, such as `?=`, `?!`, `?<=`, and `?<!`

To create effective regex patterns for allow lists, note the following tips and recommendations:

- **Anchors** – Use anchors (^ or $) only if you expect the pattern to appear at the beginning or end of a file, not the beginning or end of a line.
- **Bounded repeats** – For performance reasons, Macie limits the size of bounded repeat groups. For example, `\d{100,1000}` won't compile in Macie. To approximate this functionality, you can use an open-ended repeat such as `\d{100,}`.
- **Case insensitivity** – To make parts of a pattern case insensitive, you can use the `(?i)` construct instead of the `/i` flag.
- **Performance** – There's no need to optimize prefixes or alternations manually. For example, changing `/hello|hi|hey/` to `/h(?:ello|i|ey)/` won't improve performance.
- **Wildcards** – For performance reasons, Macie limits the number of repeated wildcards. For example, `a*b*a*` won't compile in Macie.
- **Alternation** – To specify more than one pattern in a single allow list, you can use the alternation operator (`|`) to concatenate the patterns. If you do this, Macie uses OR logic to combine the patterns and form a new pattern. For example, if you specify `(apple|orange)`, Macie recognizes both *apple* and *orange* as a match and ignores occurrences of both words. If you

concatenate patterns, be sure to limit the overall length of the concatenated expression to 512 or fewer characters.

Finally, when you develop the regex, design it to accommodate different file types and storage formats. Macie uses the regex as part of a larger data extraction and analysis framework. The framework factors an S3 object's file type or storage format. For structured data, such as columnar data in a CSV file or record-based data in a JSON file, Macie ignores text that completely matches the pattern only if all the text is stored in a single field, cell, or array. This requirement doesn't apply to structured data that's stored in an otherwise unstructured file, such as a table in an Adobe Portable Document Format (.pdf) file. For unstructured data, such as text in a .pdf file, Macie ignores text that completely matches the pattern, including text that spans multiple lines or pages.

Examples

The following examples demonstrate valid regex patterns for some common scenarios.

Email addresses

If you use a custom data identifier to detect email addresses, you can ignore email addresses that you don't consider sensitive, such as email addresses for your organization.

To ignore email addresses for a particular second-level and top-level domain, you can use this pattern:

```
[a-zA-Z0-9_.+\\-]+@example\.com
```

Where *example* is the name of the second-level domain and *com* is the top-level domain. In this case, Macie matches and ignores addresses such as *johndoe@example.com* and *john.doe@example.com*.

To ignore email addresses for a particular domain in any generic top-level domain (gTLD), such as *.com* or *.gov*, you can use this pattern:

```
[a-zA-Z0-9_.+\\-]+@example\.[a-zA-Z]{2,}
```

Where *example* is the name of the domain. In this case, Macie matches and ignores addresses such as *johndoe@example.com*, *john.doe@example.gov*, and *johndoe@example.edu*.

To ignore email addresses for a particular domain in any one country code top-level domain (ccTLD), such as *.ca* for Canada or *.au* for Australia, you can use this pattern:

```
[a-zA-Z0-9_.+\\-]+@example\.(ca|au)
```

Where *example* is the name of the domain and *ca* and *au* are specific ccTLDs to ignore. In this case, Macie matches and ignores addresses such as *johndoe@example.ca* and *john.doe@example.au*.

To ignore email addresses that are for a particular domain and gTLD and include third- and fourth-level domains, you can use this pattern:

```
[a-zA-Z0-9_.+\\-]+@([a-zA-Z0-9-]+\.)?[a-zA-Z0-9-]+\.example\.com
```

Where *example* is the name of the domain and *com* is the gTLD. In this case, Macie matches and ignores addresses such as *johndoe@www.example.com* and *john.doe@www.team.example.com*.

Phone numbers

Macie provides managed data identifiers that can detect phone numbers for several countries and regions. To ignore certain phone numbers, such as toll-free numbers or public phone numbers for your organization, you can use patterns such as the following.

To ignore toll-free, US phone numbers that use the *800* area code and are formatted as *(800) ###-####*:

```
^\(?800\)?[ -]?\d{3}[ -]?\d{4}$
```

To ignore toll-free, US phone numbers that use the *888* area code and are formatted as *(888) ###-####*:

```
^\(?888\)?[ -]?\d{3}[ -]?\d{4}$
```

To ignore 10-digit, French phone numbers that include the *33* country code and are formatted as *+33 ## ## ## ## ##*:

```
^\+33 \d( \d\d){4}$
```

To ignore US and Canadian phone numbers that use particular area and exchange codes, don't include a country code, and are formatted as *(###) ###-####*:

```
^\(?123\)?[ -]?555[ -]?\d{4}$
```

Where *123* is the area code and *555* is the exchange code.

To ignore US and Canadian phone numbers that use particular area and exchange codes, include a country code, and are formatted as *+1 (###) ###-####*:

```
^\+1\(?123\)?[ -]?555[ -]?\d{4}$
```

Where *123* is the area code and *555* is the exchange code.

Creating an allow list

In Amazon Macie, an allow list defines specific text or a text pattern that you want Macie to ignore when it inspects Amazon Simple Storage Service (Amazon S3) objects for sensitive data. If text matches an entry or pattern in an allow list, Macie doesn't report the text in sensitive data findings, statistics, or other types of results. This is the case even if the text matches the criteria of a managed data identifier or a custom data identifier.

You can create the following types of allow lists in Macie.

Predefined text

Use this type of list to specify words, phrases, and other kinds of character sequences that aren't sensitive, aren't likely to change, and don't necessarily adhere to a common pattern. Examples are: the names of public representatives for your organization, specific phone numbers, and specific sample data that your organization uses for testing. If you use this type of list, Macie ignores text that exactly matches an entry in the list.

For this type of list, you create a line-delimited plaintext file that lists specific text to ignore. You then store the file in an S3 bucket and configure settings for Macie to access the list in the bucket. You can then create and configure sensitive data discovery jobs to use the list, or add the list to your settings for automated sensitive data discovery. When each job starts to run or the next automated discovery analysis cycle starts, Macie retrieves the latest version of the list from Amazon S3. Macie then uses that version of the list when it inspects S3 objects for sensitive data. If Macie finds text that exactly matches an entry in the list, Macie doesn't report that occurrence of text as sensitive data.

Regular expression

Use this type of list to specify a regular expression (*regex*) that defines a text pattern to ignore. Examples are: public phone numbers for your organization, email addresses for your organization's domain, and patterned sample data that your organization uses for testing. If

you use this type of list, Macie ignores text that completely matches the regex pattern defined by the list.

For this type of list, you create a regex that defines a common pattern for text that isn't sensitive but varies or is likely to change. Unlike a list of predefined text, you create and store the regex and all other list settings in Macie. You can then create and configure sensitive data discovery jobs to use the list, or add the list to your settings for automated sensitive data discovery. When those jobs run or Macie performs automated discovery, Macie uses the latest version of the list's regex to analyze data. If Macie finds text that completely matches the pattern defined by the list, Macie doesn't report that occurrence of text as sensitive data.

For detailed requirements, recommendations, and examples of each type, see Configuration options and requirements for allow lists.

You can create as many as 10 allow lists in each supported AWS Region: up to five allow lists that specify predefined text, and up to five allow lists that specify regular expressions. You can create and use allow lists in all the AWS Regions where Macie is currently available except the Asia Pacific (Osaka) Region.

To create an allow list

How you create an allow list depends on the type of list that you want to create: a file that lists predefined text to ignore, or a regular expression that defines a text pattern to ignore. The following sections provide instructions for each type. Choose the section for the type of list that you want to create.

Predefined text

Before you create this type of allow list in Macie, do the following:

1. By using a text editor, create a line-delimited plaintext file that lists specific text to ignore—for example, a .txt, .text, or .plain file. For more information, see Syntax requirements.
2. Upload the file to an S3 general purpose bucket and note the name of the bucket and the object. You'll need to enter these names when you configure the settings in Macie.
3. Ensure that the settings for the S3 bucket and object allow you and Macie to retrieve the list from the bucket. For more information, see Storage requirements.
4. If you encrypted the S3 object, ensure that it's encrypted with a key that you and Macie are allowed to use. For more information, see Encryption/Decryption requirements.

After you complete these tasks, you're ready to configure the list's settings in Macie. You can configure the settings by using the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to configure the settings for an allow list by using the Amazon Macie console.

To configure allow list settings in Macie

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, under **Settings**, choose **Allow lists**.
3. On the **Allow lists** page, choose **Create**.
4. Under **Select a list type**, choose **Predefined text**.
5. Under **List settings**, use the following options to enter additional settings for the allow list:
   - For **Name**, enter a name for the list. The name can contain as many as 128 characters.
   - For **Description**, optionally enter a brief description of the list. The description can contain as many as 512 characters.
   - For **S3 bucket name**, enter the name of the bucket that stores the list.

     In Amazon S3, you can find this value in the **Name** field of the bucket's properties. This value is case sensitive. In addition, don't use wildcard characters or partial values when you enter the name.
   - For **S3 object name**, enter the name of the S3 object that stores the list.

     In Amazon S3, you can find this value in the **Key** field of the object's properties. If the name includes a path, be sure to include the complete path when you enter the name, for example **`allowlists/macie/mylist.txt`**. This value is case sensitive. In addition, don't use wildcard characters or partial values when you enter the name.
6. (Optional) Under **Tags**, choose **Add tag**, and then enter as many as 50 tags to assign to the allow list.

   A *tag* is a label that you define and assign to certain types of AWS resources. Each tag consists of a required tag key and an optional tag value. Tags can help you identify, categorize, and manage resources in different ways, such as by purpose, owner, environment, or other criteria. To learn more, see Tagging Macie resources.

7. When you finish, choose **Create**.

Macie tests the list's settings. Macie also verifies that it can retrieve the list from Amazon S3 and parse the list's content. If an error occurs, Macie displays a message that describes the error. For detailed information that can help you troubleshoot the error, see Options and requirements for lists of predefined text. After you address any errors, you can save the list's settings.

API

To configure allow list settings programmatically, use the CreateAllowList operation of the Amazon Macie API and specify the appropriate values for the required parameters.

For the `criteria` parameter, use an `s3WordsList` object to specify the name of the S3 bucket (`bucketName`) and the name of the S3 object (`objectKey`) that stores the list. To determine the bucket name, refer to the `Name` field in Amazon S3. To determine the object name, refer to the `Key` field in Amazon S3. Note that these values are case sensitive. In addition, don't use wildcard characters or partial values when you specify these names.

To configure the settings by using the AWS CLI, run the create-allow-list command and specify the appropriate values for the required parameters. The following examples show how to configure the settings for an allow list that's stored in an S3 bucket named *amzn-s3-demo-bucket*. The name of the S3 object that stores the list is *allowlists/macie/mylist.txt*.

This example is formatted for Linux, macOS, or Unix, and it uses the backslash (\) line-continuation character to improve readability.

```
$ aws macie2 create-allow-list \
--criteria '{"s3WordsList":{"bucketName":"amzn-s3-demo-
bucket","objectKey":"allowlists/macie/mylist.txt"}}' \
--name my_allow_list \
--description "Lists public phone numbers and names for Example Corp."
```

This example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws macie2 create-allow-list ^
--criteria={\"s3WordsList\":{\"bucketName\":\"amzn-s3-demo-bucket\",\"objectKey\":
\"allowlists/macie/mylist.txt\"}} ^
--name my_allow_list ^
--description "Lists public phone numbers and names for Example Corp."
```

When you submit your request, Macie tests the list's settings. Macie also verifies that it can retrieve the list from Amazon S3 and parse the list's content. If an error occurs, your request fails and Macie returns a message that describes the error. For detailed information that can help you troubleshoot the error, see Options and requirements for lists of predefined text.

If Macie can retrieve and parse the list, your request succeeds and you receive output similar to the following.

```
{
    "arn": "arn:aws:macie2:us-west-2:123456789012:allow-list/
nkr81bmtu2542yyexample",
    "id": "nkr81bmtu2542yyexample"
}
```

Where `arn` is the Amazon Resource Name (ARN) of the allow list that was created, and `id` is the unique identifier for the list.

After you save the list's settings, you can create and configure sensitive data discovery jobs to use the list, or add the list to your settings for automated sensitive data discovery. Each time those jobs start to run or an automated discovery analysis cycle starts, Macie retrieves the latest version of the list from Amazon S3. Macie then uses that version of the list when it analyzes data.

Regular expression

When you create an allow list that specifies a regular expression (*regex*), you define the regex and all other list settings directly in Macie. For the regex, Macie supports a subset of the pattern syntax provided by the Perl Compatible Regular Expressions (PCRE) library. For more information, see Syntax support and recommendations.

You can create this type of list by using the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to create an allow list by using the Amazon Macie console.

To create an allow list by using the console

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, under **Settings**, choose **Allow lists**.
3. On the **Allow lists** page, choose **Create**.

4. Under **Select a list type**, choose **Regular expression**.
5. Under **List settings**, use the following options to enter additional settings for the allow list:
   - For **Name**, enter a name for the list. The name can contain as many as 128 characters.
   - For **Description**, optionally enter a brief description of the list. The description can contain as many as 512 characters.
   - For **Regular expression**, enter the regex that defines the text pattern to ignore. The regex can contain as many as 512 characters.
6. (Optional) For **Evaluate**, enter up to 1,000 characters in the **Sample data** box, and then choose **Test** to test the regex. Macie evaluates the sample data and reports the number of occurrences of text that match the regex. You can repeat this step as many times as you like to refine and optimize the regex.

   > **Note**
   >
   > We recommend that you test and refine the regex with multiple sets of sample data. If you create a regex that's too general, Macie might ignore occurrences of text that you consider sensitive. If a regex is too specific, Macie might not ignore occurrences of text that you don't consider sensitive.

7. (Optional) Under **Tags**, choose **Add tag**, and then enter as many as 50 tags to assign to the allow list.

   A *tag* is a label that you define and assign to certain types of AWS resources. Each tag consists of a required tag key and an optional tag value. Tags can help you identify, categorize, and manage resources in different ways, such as by purpose, owner, environment, or other criteria. To learn more, see Tagging Macie resources.
8. When you finish, choose **Create**.

Macie tests the list's settings. Macie also tests the regex to verify that it can compile the expression. If an error occurs, Macie displays a message that describes the error. For detailed information that can help you troubleshoot the error, see Options and requirements for regular expressions. After you address any errors, you can save the allow list.

API

Before you create this type of allow list in Macie, we recommend that you test and refine the regex with multiple sets of sample data. If you create a regex that's too general, Macie might

ignore occurrences of text that you consider sensitive. If a regex is too specific, Macie might not ignore occurrences of text that you don't consider sensitive.

To test an expression with Macie, you can use the TestCustomDataIdentifier operation of the Amazon Macie API or, for the AWS CLI, run the test-custom-data-identifier command. Macie uses the same underlying code to compile expressions for allow lists and custom data identifiers. If you test an expression in this way, be sure to specify values only for the `regex` and `sampleText` parameters. Otherwise, you'll receive inaccurate results.

When you're ready to create this type of allow list, use the CreateAllowList operation of the Amazon Macie API and specify the appropriate values for the required parameters. For the `criteria` parameter, use the `regex` field to specify the regular expression that defines the text pattern to ignore. The expression can contain as many as 512 characters.

To create this type of list by using the AWS CLI, run the create-allow-list command and specify the appropriate values for the required parameters. The following examples create an allow list named *`my_allow_list`*. The regex is designed to ignore all email addresses that a custom data identifier might otherwise detect for the `example.com` domain.

This example is formatted for Linux, macOS, or Unix, and it uses the backslash (\) line-continuation character to improve readability.

```
$ aws macie2 create-allow-list \
--criteria '{"regex":"[a-z]@example.com"}' \
--name my_allow_list \
--description "Ignores all email addresses for Example Corp."
```

This example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws macie2 create-allow-list ^
--criteria={\"regex\":\"[a-z]@example.com\"} ^
--name my_allow_list ^
--description "Ignores all email addresses for Example Corp."
```

When you submit your request, Macie tests the list's settings. Macie also tests the regex to verify that it can compile the expression. If an error occurs, the request fails and Macie returns a message that describes the error. For detailed information that can help you troubleshoot the error, see Options and requirements for regular expressions.

If Macie can compile the expression, the request succeeds and you receive output similar to the following:

```
{
    "arn": "arn:aws:macie2:us-west-2:123456789012:allow-list/
km2d4y22hp6rv05example",
    "id": "km2d4y22hp6rv05example"
}
```

Where `arn` is the Amazon Resource Name (ARN) of the allow list that was created, and `id` is the unique identifier for the list.

After you save the list, you can create and configure sensitive data discovery jobs to use it, or add it to your settings for automated sensitive data discovery. When those jobs run or Macie performs automated discovery, Macie uses the latest version of the list's regex to analyze data.

Checking the status of an allow list

If you create an allow list, it's important to check its status periodically. Otherwise, errors might cause Amazon Macie to produce unexpected analysis results for your Amazon Simple Storage Service (Amazon S3) data. For example, Macie might create sensitive data findings for text that you specified in an allow list.

If you configure a sensitive data discovery job to use an allow list and Macie can't access or use the list when the job starts to run, the job continues to run. However, Macie doesn't use the list when it analyzes S3 objects. Similarly, if an analysis cycle starts for automated sensitive data discovery and Macie can't access or use a specified allow list, the analysis continues but Macie doesn't use the list.

Errors are unlikely to occur for an allow list that specifies a regular expression (*regex*). This is partly because Macie automatically tests the regex when you create or update the list's settings. In addition, you store the regex and all other list settings in Macie.

However, errors can occur for an allow list that specifies predefined text, partly because you store the list in Amazon S3 instead of Macie. Common causes of errors are:

- The S3 bucket or object is deleted.
- The S3 bucket or object is renamed and the list's settings in Macie don't specify the new name.
- The S3 bucket's permissions settings are changed and Macie loses access to the bucket and the object.

- The encryption settings for the S3 bucket are changed and Macie can't decrypt the object that stores the list.
- The policy for the encryption key is changed and Macie loses access to the key. Macie can't decrypt the S3 object that stores the list.

> **Important**
>
> Because these errors affect your analyses' results, we recommend that you check the status of all of your allow lists periodically. We recommend that you also do this if you change the permissions or encryption settings for an S3 bucket that stores an allow list, or you change the policy for an AWS Key Management Service (AWS KMS) key that's used to encrypt a list.

For detailed information that can help you troubleshoot errors that occur, see Options and requirements for lists of predefined text.

To check the status of an allow list

You can check the status of an allow list by using the Amazon Macie console or the Amazon Macie API. On the console, you can use a single page to check the status of all of your allow lists at the same time. If you use the Amazon Macie API, you can check the status of individual allow lists, one at a time.

Console

Follow these steps to check the status of your allow lists by using the Amazon Macie console.

To check the status of your allow lists

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, under **Settings**, choose **Allow lists**.
3. On the **Allow lists** page, choose refresh ().
   Macie tests the settings for all of your allow lists and updates the **Status** field to indicate the current status of each list.

If a list specifies a regular expression, its status is typically **OK**. This means that Macie can compile the expression. If a list specifies predefined text, its status can be any of the following values.

OK

Macie can retrieve and parse the contents of the list.

Access denied

Macie isn't allowed to access the S3 object that stores the list. Amazon S3 denied the request to retrieve the object. A list can also have this status if the object is encrypted with a customer managed AWS KMS key that Macie isn't allowed to use.

To address this error, review the bucket policy and other permissions settings for the bucket and the object. Ensure that Macie is allowed to access and retrieve the object. If the object is encrypted with a customer managed AWS KMS key, also review the key policy and ensure that Macie is allowed to use the key.

Error

A transient or internal error occurred when Macie attempted to retrieve or parse the contents of the list. An allow list can also have this status if it's encrypted with an encryption key that Amazon S3 and Macie can't access or use.

To address this error, wait a few minutes and then choose refresh () again. If the status continues to be **Error**, check the encryption settings for the S3 object. Ensure that the object is encrypted with a key that Amazon S3 and Macie can access and use.

Object is empty

Macie can retrieve the list from Amazon S3 but the list doesn't contain any content.

To address this error, download the object from Amazon S3 and ensure that it contains the correct entries. If the entries are correct, review the list's settings in Macie. Ensure that the specified bucket and object names are correct.

Object not found

The list doesn't exist in Amazon S3.

To address this error, review the list's settings in Macie. Ensure that the specified bucket and object names are correct.

Quota exceeded

Macie can access the list in Amazon S3. However, the number of entries in the list or the storage size of the list exceeds the quota for an allow list.

To address this error, break the list into multiple files. Ensure that each file contains fewer than 100,000 entries. Also ensure that the size of each file is less than 35 MB. Then, upload each file to Amazon S3. When you finish, configure allow list settings in Macie for each file. You can have as many as five lists of predefined text in each supported AWS Region.

Throttled

Amazon S3 throttled the request to retrieve the list.

To address this error, wait a few minutes and then choose refresh
()
again.

User access denied

Amazon S3 denied the request to retrieve the object. If the specified object exists, you're not allowed to access it or it's encrypted with an AWS KMS key that you're not allowed to use.

To address this error, work with your AWS administrator to ensure that the list's settings specify the correct bucket and object names, and you have read access to the bucket and the object. If the object is encrypted, also ensure that it's encrypted with a key that you're allowed to use.

4. To review the settings and status of a specific list, choose the list's name.

API

To check the status of an allow list programmatically, use the GetAllowList operation of the Amazon Macie API. Or, if you're using the AWS CLI, run the get-allow-list command.

For the `id` parameter, specify the unique identifier for the allow list whose status you want to check. To get this identifier, you can use the ListAllowLists operation. The **ListAllowLists** operation retrieves information about all the allow lists for your account. If you're using the AWS CLI, you can run the list-allow-lists command to retrieve this information.

When you submit a **GetAllowList** request, Macie tests all the settings for the allow list. If the settings specify a regular expression (`regex`), Macie verifies that it can compile the expression. If the settings specify a list of predefined text (`s3WordsList`), Macie verifies that it can retrieve and parse the list.

Macie then returns a `GetAllowListResponse` object that provides the details of the allow list. In the `GetAllowListResponse` object, the `status` object indicates the current status of the list: a status code (`code`) and, depending on the status code, a brief description of the list's status (`description`).

If the allow list specifies a regex, the status code is typically `OK` and there isn't an associated description. This means that Macie compiled the expression successfully.

If the allow list specifies predefined text, the status code varies depending on the test results:

- If Macie retrieved and parsed the list successfully, the status code is `OK` and there isn't an associated description.
- If an error prevented Macie from retrieving or parsing the list, the status code and description indicate the nature of the error that occurred.

For a list of possible status codes and a description of each one, see AllowListStatus in the *Amazon Macie API Reference*.

Changing an allow list

After you create an allow list, you can change most of the list's settings in Amazon Macie. For example, you can change the list's name and description. You can also add and edit tags for the list. The only setting that you can't change is a list's type. For example, if an existing list specifies a regular expression (*regex*), you can't change its type to predefined text.

If an allow list specifies predefined text, you can also change the entries in the list. To do this, update the file that contains the entries. Then upload the new version of the file to Amazon Simple Storage Service (Amazon S3). The next time Macie prepares to use the list, Macie retrieves the latest version of the file from Amazon S3. When you upload the new file, ensure that you store it in the same S3 bucket and object. Or, if you change the name of the bucket or object, ensure that you update the list's settings in Macie.

To change the settings for an allow list

You can change the settings for an allow list by using the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to change an allow list's settings by using the Amazon Macie console.

To change an allow list's settings by using the console

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, under **Settings**, choose **Allow lists**.
3. On the **Allow lists** page, choose the name of the allow list that you want to change. The allow list page opens and displays the current settings for the list.
4. To add or edit tags for the allow list, choose **Manage tags** in the **Tags** section. Then change the tags as necessary. When you finish, choose **Save**.
5. To change other settings for the allow list, choose **Edit** in the **List settings** section. Then change the settings that you want:

   - **Name** – Enter a new name for the list. The name can contain as many as 128 characters.
   - **Description** – Enter a new description of the list. The description can contain as many as 512 characters.
   - If the allow list specifies predefined text:
     - **S3 bucket name** – Enter the name of the bucket that stores the list.

       In Amazon S3, you can find this value in the **Name** field of the bucket's properties. This value is case sensitive. In addition, don't use wildcard characters or partial values when you enter the name.
     - **S3 object name** – Enter the name of the S3 object that stores the list.

In Amazon S3, you can find this value in the **Key** field of the object's properties. If the name includes a path, be sure to include the complete path when you enter the name, for example **`allowlists/macie/mylist.txt`**. This value is case sensitive. In addition, don't use wildcard characters or partial values when you enter the name.

- If the allow list specifies a regular expression (*regex*), enter a new regex in the **Regular expression** box. The regex can contain as many as 512 characters.

  After you enter the new regex, optionally test it. To do this, enter up to 1,000 characters in the **Sample data** box, and then choose **Test**. Macie evaluates the sample data and reports the number of occurrences of text that match the regex. You can repeat this step as many times as you like to refine and optimize the regex before you save your changes.

6. When you finish, choose **Save**.

Macie tests the list's settings. For a list of predefined text, Macie also verifies that it can retrieve the list from Amazon S3 and parse the list's content. For a regex, Macie also verifies that it can compile the expression. If an error occurs, Macie displays a message that describes the error. For detailed information that can help you troubleshoot the error, see Configuration options and requirements for allow lists. After you address any errors, you can save your changes.

API

To change an allow list's settings programmatically, use the UpdateAllowList operation of the Amazon Macie API. Or, if you're using the AWS CLI, run the update-allow-list command. In your request, use the supported parameters to specify a new value for each setting that you want to change. Note that the `criteria`, `id`, and `name` parameters are required. If you don't want to change the value for a required parameter, specify the current value for the parameter.

For example, the following command changes the name and description of an existing allow list. The example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws macie2 update-allow-list ^
--id km2d4y22hp6rv05example ^
--name my_allow_list-email ^
--criteria={\"regex\":\"[a-z]@example.com\"} ^
--description "Ignores all email addresses for the example.com domain"
```

Where:

- *km2d4y22hp6rv05example* is the unique identifier for the list.
- *my_allow_list-email* is the new name for the list.
- *[a-z]@example.com* is the list's criteria, a regular expression.
- *Ignores all email addresses for the example.com domain* is the new description for the list.

When you submit your request, Macie tests the list's settings. If the list specifies predefined text (`s3WordsList`), this includes verifying that Macie can retrieve the list from Amazon S3 and parse the list's content. If the list specifies a regex (`regex`), this includes verifying that Macie can compile the expression.

If an error occurs when Macie tests the settings, your request fails and Macie returns a message that describes the error. For detailed information that can help you troubleshoot the error, see Configuration options and requirements for allow lists. If the request fails for another reason, Macie returns an HTTP 4*xx* or 500 response that indicates why the operation failed.

If your request succeeds, Macie updates the list's settings and you receive output similar to the following.

```
{
    "arn": "arn:aws:macie2:us-west-2:123456789012:allow-list/
km2d4y22hp6rv05example",
    "id": "km2d4y22hp6rv05example"
}
```

Where `arn` is the Amazon Resource Name (ARN) of the allow list that was updated, and `id` is the unique identifier for the list.

Deleting an allow list

When you delete an allow list in Amazon Macie, you permanently delete all the list's settings. These settings can't be recovered after they're deleted. If the settings specify a list of predefined text that you store in Amazon Simple Storage Service (Amazon S3), Macie doesn't delete the S3 object that stores the list. Only the settings in Macie are deleted.

If you configure sensitive data discovery jobs to use an allow list that you subsequently delete, the jobs will run as scheduled. However, your job results, both sensitive data findings and sensitive

data discovery results, might report text that you previously specified in the allow list. Similarly, if you configure automated sensitive data discovery to use a list that you subsequently delete, daily analyses cycles will proceed. However, sensitive data findings, statistics, and other types of results might report text that you previously specified in the allow list.

Before you delete an allow list, we recommend that you review your job inventory to identify jobs that use the list and are scheduled to run in the future. In the inventory, the details panel indicates whether a job is configured to use any allow lists and, if so, which ones. We recommend that you also check your settings for automated sensitive data discovery. You might determine that it's best to change a list instead of deleting it.

As an additional safeguard, Macie checks the settings for all of your jobs when you try to delete an allow list. If you configured jobs to use the list and any of those jobs have a status other than **Complete** or **Cancelled**, Macie doesn't delete the list unless you provide additional confirmation.

To delete an allow list

You can delete an allow list by using the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to delete an allow list by using the Amazon Macie console.

To delete an allow list by using the console

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, under **Settings**, choose **Allow lists**.
3. On the **Allow lists** page, select the checkbox for the allow list that you want to delete.
4. On the **Actions** menu, choose **Delete**.
5. When prompted for confirmation, enter **`delete`**, and then choose **Delete**.

API

To delete an allow list programmatically, use the DeleteAllowList operation of the Amazon Macie API. For the `id` parameter, specify the unique identifier for the allow list to delete. You can get this identifier by using the ListAllowLists operation. The **ListAllowLists** operation retrieves information about all the allow lists for your account. If you're using the AWS CLI, you can run the list-allow-lists command to retrieve this information.

For the `ignoreJobChecks` parameter, specify whether to force deletion of the list, even if sensitive data discovery jobs are configured to use the list:

- If you specify `false`, Macie checks the settings for all of your jobs that have a status other than `COMPLETE` or `CANCELLED`. If none of those jobs are configured to use the list, Macie deletes the list permanently. If any of those jobs are configured to use the list, Macie rejects your request and returns an HTTP 400 (`ValidationException`) error. The error message indicates the number of applicable jobs for up to 200 jobs.
- If you specify `true`, Macie deletes the list permanently without checking the settings for any of your jobs.

To delete an allow list by using the AWS CLI, run the delete-allow-list command. For example:

```
C:\> aws macie2 delete-allow-list --id nkr81bmtu2542yyexample --ignore-job-checks
 false
```

Where *nkr81bmtu2542yyexample* is the unique identifier for the allow list to delete.

If your request succeeds, Macie returns an empty HTTP 200 response. Otherwise, Macie returns an HTTP 4*xx* or 500 response that indicates why the operation failed.

If the allow list specified predefined text, you can optionally delete the S3 object that stores the list. However, keeping this object can help ensure that you have an immutable history of sensitive data findings and discovery results for data privacy and protection audits or investigations.

Performing automated sensitive data discovery

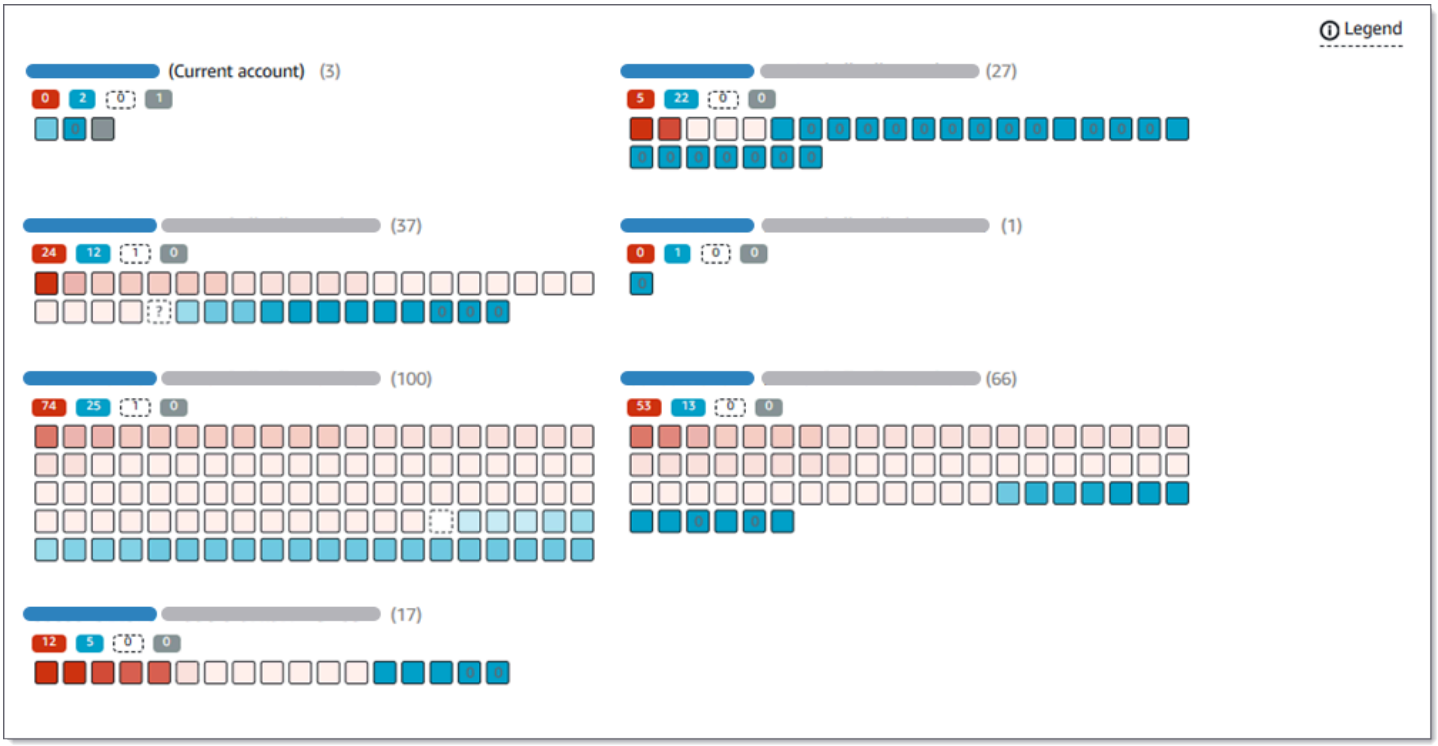
For broad visibility into where sensitive data might reside in your Amazon Simple Storage Service (Amazon S3) data estate, configure Amazon Macie to perform automated sensitive data discovery for your account or organization. With automated sensitive data discovery, Macie continually evaluates your S3 bucket inventory and uses sampling techniques to identify and select representative S3 objects in your buckets. Macie then retrieves and analyzes the selected objects, inspecting them for sensitive data.

By default, Macie selects and analyzes objects from all of your S3 general purpose buckets. If you're the Macie administrator for an organization, this includes objects in buckets that your member accounts own. You can adjust the scope of the analyses by excluding specific buckets. For example,

you might exclude buckets that typically store AWS logging data. If you're a Macie administrator, an additional option is to enable or disable automated sensitive data discovery for individual accounts in your organization on a case-by-case basis.

You can tailor the analyses to focus on specific types of sensitive data. By default, Macie analyzes S3 objects by using the set of managed data identifiers that we recommend for automated sensitive data discovery. To tailor the analyses, you can configure Macie to use specific managed data identifiers that Macie provides, custom data identifiers that you define, or a combination of the two. You can also refine the analyses by configuring Macie to use allow lists that you specify.

As the analysis progresses each day, Macie produces records of the sensitive data that it finds and the analysis that it performs: *sensitive data findings*, which report sensitive data that Macie finds in individual S3 objects, and *sensitive data discovery results*, which log details about the analysis of individual S3 objects. Macie also updates statistics, inventory data, and other information that it provides about your Amazon S3 data. For example, an interactive heat map on the console provides a visual representation of data sensitivity across your data estate:



These features are designed to help you evaluate data sensitivity across your Amazon S3 data estate, and drill down to investigate and assess individual accounts, buckets, and objects. They can also help you determine where to perform deeper, more immediate analysis by running sensitive data discovery jobs. Combined with information that Macie provides about the security and privacy of your Amazon S3 data, you can also use these features to identify cases where immediate

remediation might be necessary—for example, a publicly accessible bucket that Macie found sensitive data in.

To configure and manage automated sensitive data discovery, you must be the Macie administrator for an organization or have a standalone Macie account.

Topics

How automated sensitive data discovery works

When you enable Amazon Macie for your AWS account, Macie creates an AWS Identity and Access Management (IAM) service-linked role for your account in the current AWS Region. The permissions policy for this role allows Macie to call other AWS services and monitor AWS resources on your behalf. By using this role, Macie generates and maintains an inventory of your Amazon Simple Storage Service (Amazon S3) general purpose buckets in the Region. The inventory includes information about each of your S3 buckets and objects in the buckets. If you're the Macie administrator for an organization, your inventory includes information about buckets that your member accounts own. For more information, see Managing multiple accounts.

If you enable automated sensitive data discovery, Macie evaluates your inventory data on a daily basis to identify S3 objects that are eligible for automated discovery. As part of the evaluation, Macie also selects a sampling of representative objects to analyze. Macie then retrieves and analyzes the latest version of each selected object, inspecting it for sensitive data.

As the analysis progresses each day, Macie updates statistics, inventory data, and other information that it provides about your Amazon S3 data. Macie also produces records of the sensitive data it finds and the analysis that it performs. The resulting data provides insight into where Macie found sensitive data in your Amazon S3 data estate, which can span all the S3 general purpose buckets for your account. The data can help you assess the security and privacy of your Amazon S3

data, determine where to perform a deeper investigation, and identify cases where remediation is necessary.

For a brief demonstration of how automated sensitive data discovery works, watch the following video: Amazon Macie automated data discovery overview.

To configure and manage automated sensitive data discovery, you must be the Macie administrator for an organization or have a standalone Macie account. If your account is part of an organization, only the Macie administrator for your organization can enable or disable automated discovery for accounts in the organization. In addition, only the Macie administrator can configure and manage automated discovery settings for the accounts. This includes settings that define the scope and nature of the analyses that Macie performs. If you have a member account in an organization, contact your Macie administrator to learn about the settings for your account and organization.

Topics

Key components

Amazon Macie uses a combination of features and techniques to perform automated sensitive data discovery. These work together with features that Macie provides to help you monitor your Amazon S3 data for security and access control.

Selecting S3 objects to analyze

On a daily basis, Macie evaluates your Amazon S3 inventory data to identify S3 objects that are eligible for analysis by automated sensitive data discovery. If you're the Macie administrator for an organization, by default the evaluation includes data for S3 buckets that your member accounts own.

As part of the evaluation, Macie uses sampling techniques to select representative S3 objects to analyze. The techniques define groups of objects that have similar metadata and are likely to have similar content. The groups are based on dimensions such as bucket name, prefix, storage class, file name extension, and last modified date. Macie then selects a representative set of samples from each group, retrieves the latest version of each selected object from Amazon S3, and analyzes each selected object to determine whether the object contains sensitive data. When the analysis is complete, Macie discards its copy of the object.

The sampling strategy prioritizes distributed analyses. In general, it uses a breadth-first approach to your Amazon S3 data estate. Each day, a representative set of S3 objects are selected from as many of your general purpose buckets as possible based on the total storage size of all the classifiable objects in your Amazon S3 data estate. For example, if Macie has already analyzed and found sensitive data in objects in one bucket and hasn't yet analyzed objects in another bucket, the latter bucket is a higher priority for analysis. With this approach, you gain broad insight into the sensitivity of your Amazon S3 data more quickly. Depending on the size of your data estate, analysis results can begin to appear within 48 hours.

The sampling strategy also prioritizes analysis of different kinds of S3 objects and objects that were recently created or changed. Any single object sample isn’t guaranteed to be conclusive. Therefore, analysis of a diverse set of objects can yield better insight into the types and amount of sensitive data that an S3 bucket might contain. In addition, prioritizing new or recently changed objects helps the analysis adapt to changes to your bucket inventory. For example, if objects are created or changed after a previous analysis, those objects are a higher priority for subsequent analysis. Conversely, if an object was previously analyzed and hasn't changed since that analysis, Macie doesn't analyze the object again. This approach helps you establish sensitivity baselines for individual S3 buckets. Then, as continual, incremental analyses progress for your account, your sensitivity assessments of individual buckets can become increasingly deeper and detailed at a predictable rate.

Defining the scope of the analyses

By default, Macie includes all the S3 general purpose buckets for your account when it evaluates your inventory data and selects S3 objects to analyze. If you're the Macie administrator for an organization, this includes buckets that your member accounts own.

You can adjust the scope of the analyses by excluding specific S3 buckets from automated sensitive data discovery. For example, you might want to exclude buckets that typically store AWS logging data, such as AWS CloudTrail event logs. To exclude a bucket, you can change the automated discovery settings for your account or the bucket. If you do this, Macie starts excluding the bucket when the next daily evaluation and analysis cycle starts. You can exclude as many as 1,000 buckets from analyses. If you exclude an S3 bucket, you can include it again later. To do this, change the settings for your account or the bucket again. Macie then starts including the bucket when the next daily evaluation and analysis cycle starts.

If you're the Macie administrator for an organization, you can also enable or disable automated sensitive data discovery for individual accounts in your organization. If you disable automated discovery for an account, Macie excludes all the S3 buckets that the account owns. If you

subsequently re-enable automated discovery for the account, Macie starts including the buckets again.

Determining which types of sensitive data to detect and report

By default, Macie inspects S3 objects by using the set of managed data identifiers that we recommend for automated sensitive data discovery. For a list of these managed data identifiers, see Default settings for automated sensitive data discovery.

You can tailor the analyses to focus on specific types of sensitive data. To do this, change your automated discovery settings in any of the following ways:

- **Add or remove managed data identifiers** – A *managed data identifier* is a set of built-in criteria and techniques that are designed to detect a specific type of sensitive data, such as credit card numbers, AWS secret access keys, or passport numbers for a particular country or region. For more information, see Using managed data identifiers.
- **Add or remove custom data identifiers** – A *custom data identifier* is a set of criteria that you define to detect sensitive data. With custom data identifiers, you can detect sensitive data that reflects your organization's particular scenarios, intellectual property, or proprietary data. For example, you can detect employee IDs, customer account numbers, or internal data classifications. For more information, see Building custom data identifiers.
- **Add or remove allow lists** – In Macie, an allow list specifies text or a text pattern that you want Macie to ignore in S3 objects. These are typically sensitive data exceptions for your particular scenarios or environment, such as public names or phone numbers for your organization, or sample data that your organization uses for testing. For more information, see Defining sensitive data exceptions with allow lists.

If you change a setting, Macie applies your change when the next daily analysis cycle starts. If you're the Macie administrator for an organization, Macie uses the settings for your account when it analyzes S3 objects for other accounts in your organization.

You can also configure bucket-level settings that determine whether specific types of sensitive data are included in assessments of a bucket's sensitivity. To learn how, see Adjusting sensitivity scores for S3 buckets.

Calculating sensitivity scores

By default, Macie automatically calculates a sensitivity score for each S3 general purpose bucket for your account. If you're the Macie administrator for an organization, this includes buckets that your member accounts own.

In Macie, a *sensitivity score* is a quantitative measure of the intersection of two primary dimensions: the amount of sensitive data that Macie has found in a bucket, and the amount of data that Macie has analyzed in a bucket. A bucket's sensitivity score determines which sensitivity label Macie assigns to the bucket. A *sensitivity label* is a qualitative representation of a bucket's sensitivity score—for example, *Sensitive*, *Not sensitive*, and *Not yet analyzed*. For details about the range of sensitivity scores and labels that Macie defines, see Sensitivity scoring for S3 buckets.

> **Important**
>
> An S3 bucket's sensitivity score and label don't imply or otherwise indicate the criticality or importance that the bucket or the bucket's objects might have for you or your organization. Instead, they're intended to provide reference points that can help you identify and monitor potential security risks.

When you enable automated sensitive data discovery for the first time, Macie automatically assigns a sensitivity score of *50* and the *Not yet analyzed* label to each S3 bucket. The exception is empty buckets. An *empty bucket* is a bucket that doesn't store any objects or all the bucket's objects contain zero (0) bytes of data. If this is the case for a bucket, Macie assigns a score of *1* to the bucket and it assigns the *Not sensitive* label to the bucket.

As automated sensitive data discovery progresses, Macie updates sensitivity scores and labels to reflect the results of its analyses. For example:

- If Macie doesn't find sensitive data in an object, Macie decreases the bucket's sensitivity score and updates the bucket's sensitivity label as necessary.
- If Macie finds sensitive data in an object, Macie increases the bucket's sensitivity score and updates the bucket's sensitivity label as necessary.
- If Macie finds sensitive data in an object that's subsequently changed, Macie removes sensitive data detections for the object from the bucket's sensitivity score and updates the bucket's sensitivity label as necessary.
- If Macie finds sensitive data in an object that's subsequently deleted, Macie removes sensitive data detections for the object from the bucket's sensitivity score and updates the bucket's sensitivity label as necessary.

You can adjust the sensitivity scoring settings for individual S3 buckets by including or excluding specific types of sensitive data from a bucket's score. You can also override a bucket's

calculated score by manually assigning the maximum score (*100*) to the bucket. If you assign the maximum score, the bucket's label is *Sensitive*. For more information, see Adjusting sensitivity scores for S3 buckets.

Generating metadata, statistics, and other types of results

When you enable automated sensitive data discovery, Macie generates and begins maintaining additional inventory data, statistics, and other information about the S3 general purpose buckets for your account. If you're the Macie administrator for an organization, by default this includes buckets that your member accounts own.

The additional information captures the results of the automated sensitive data discovery activities that Macie has performed thus far. It also supplements other information that Macie provides about your Amazon S3 data, such as the public access and shared access settings for individual buckets. The additional information includes:

- An interactive, visual representation of data sensitivity across your Amazon S3 data estate.
- Aggregated data sensitivity statistics, such as the total number of buckets that Macie has found sensitive data in and how many of those buckets are publicly accessible.
- Bucket-level details that indicate the current status of the analyses. For example, a list of objects that Macie has analyzed in a bucket, the types of sensitive data that Macie has found in a bucket, and the number of occurrences of each type of sensitive data that Macie found.

The information also includes statistics and details that can help you assess and monitor coverage of your Amazon S3 data. You can check the status of the analyses for your data estate overall and for individual S3 buckets. You can also identify issues that prevented Macie from analyzing objects in specific buckets. If you remediate the issues, you can increase coverage of your Amazon S3 data during subsequent analysis cycles. For more information, see Assessing automated sensitive data discovery coverage.

Macie automatically recalculates and updates this information while it performs automated sensitive data discovery. For example, if Macie finds sensitive data in an S3 object that's subsequently changed or deleted, Macie updates the applicable bucket's metadata: removes the object from the list of analyzed objects; removes occurrences of sensitive data that Macie found in the object; recalculates the sensitivity score, if the score is calculated automatically; and, updates the sensitivity label as necessary to reflect the new score.

In addition to metadata and statistics, Macie produces records of the sensitive data it finds and the analysis that it performs: *sensitive data findings*, which report sensitive data that Macie finds

in individual S3 objects, and *sensitive data discovery results*, which log details about the analysis of individual S3 objects.

For more information, see Reviewing automated sensitive data discovery results.

Considerations

As you configure and use Amazon Macie to perform automated sensitive data discovery for your Amazon S3 data, keep the following in mind:

- Your automated discovery settings apply only to the current AWS Region. Consequently, the resulting analyses and data apply only to S3 general purpose buckets and objects in the current Region. To perform automated discovery and access the resulting data in additional Regions, enable and configure automated discovery in each additional Region.
- If you're the Macie administrator for an organization:
  - You can perform automated discovery for a member account only if Macie is enabled for the account in the current Region. In addition, you must enable automated discovery for the account in that Region. Members can't enable or disable automated discovery for their own accounts.
  - If you enable automated discovery for a member account, Macie uses the automated discovery settings for your administrator account when it analyzes data for the member account. The applicable settings are: the list of S3 buckets to exclude from analyses, and the managed data identifiers, custom data identifiers, and allow lists to use when analyzing S3 objects. Members can't review or change these settings.
  - Members can't access automated discovery settings for individual S3 buckets that they own. For example, a member can't review or adjust the sensitivity scoring settings for one of their buckets. Only the Macie administrator can access these settings.
  - Members have read access to sensitive data discovery statistics and other results that Macie directly provides for their S3 buckets. For example, a member can use Macie to review sensitivity scores and coverage data for their S3 buckets. The exception is sensitive data findings. Only the Macie administrator has direct access to findings that automated discovery produces.
- If an S3 bucket's permissions settings prevent Macie from accessing or retrieving information about the bucket or the bucket's objects, Macie can't perform automated discovery for the bucket. Macie can only provide a subset of information about the bucket, such as the account ID for the AWS account that owns the bucket, the bucket's name, and when Macie most recently

retrieved bucket and object metadata for the bucket as part of the daily refresh cycle. In your bucket inventory, the sensitivity score for these buckets is *50* and their sensitivity label is *Not yet analyzed*. To identify S3 buckets where this is the case, you can refer to coverage data. For more information, see Assessing automated sensitive data discovery coverage.

- To be eligible for selection and analysis, an S3 object must be stored in a general purpose bucket and it must be *classifiable*. A *classifiable* object uses a supported Amazon S3 storage class and it has a file name extension for a supported file or storage format. For more information, see Supported storage classes and formats.
- If an S3 object is encrypted, Macie can analyze it only if it's encrypted with a key that Macie can access and is allowed to use. For more information, see Analyzing encrypted S3 objects. To identify cases where encryption settings prevented Macie from analyzing one or more objects in a bucket, you can refer to coverage data. For more information, see Assessing automated sensitive data discovery coverage.

Configuring automated sensitive data discovery

To gain broad visibility into where sensitive data might reside in your Amazon Simple Storage Service (Amazon S3) data estate, enable and configure automated sensitive data discovery for your account or organization. Amazon Macie then evaluates your S3 bucket inventory on a daily basis and uses sampling techniques to identify and select representative S3 objects from your buckets. Macie retrieves and analyzes the selected objects, inspecting them for sensitive data. If you're the Macie administrator for an organization, by default this includes objects in S3 buckets that your member accounts own.

As the analysis progresses each day, Macie produces records of the sensitive data it finds and the analysis that it performs. Macie also updates statistics, inventory data, and other information that it provides about your Amazon S3 data. The resulting data provides insight into where Macie found sensitive data in your Amazon S3 data estate, which can span all the S3 buckets for your account or organization. For more information, see How automated sensitive data discovery works.

If you have a standalone Macie account or you're the Macie administrator for an organization, you can configure and manage automated sensitive data discovery for your account or organization. This includes enabling and disabling automated discovery, and configuring settings that define the scope and nature of the analyses that Macie performs. If you have a member account in an organization, contact your Macie administrator to learn about the settings for your account and organization.

Topics

Prerequisites for configuring automated sensitive data discovery

Before you enable or configure settings for automated sensitive data discovery, complete the following tasks. This helps ensure that you have the resources and permissions that you need.

To complete these tasks, you must be the Amazon Macie administrator for an organization or have a standalone Macie account. If your account is part of an organization, only the Macie administrator for your organization can enable or disable automated sensitive data discovery for accounts in the organization. In addition, only the Macie administrator can configure automated discovery settings for the accounts.

Tasks

Step 1: Configure a repository for sensitive data discovery results

When Amazon Macie performs automated sensitive data discovery, it creates an analysis record for each Amazon Simple Storage Service (Amazon S3) object that it selects for analysis. These records, referred to as *sensitive data discovery results*, log details about the analysis of individual S3 objects. This includes objects that Macie doesn't find sensitive data in, and objects that Macie can't analyze due to errors or issues such as permissions settings. If Macie finds sensitive data in an object, the sensitive data discovery result includes information about the sensitive data that Macie found. Sensitive data discovery results provide you with analysis records that can be helpful for data privacy and protection audits or investigations.

Macie stores your sensitive data discovery results for only 90 days. To access the results and enable long-term storage and retention of them, configure Macie to store the results in an S3 bucket. The bucket can serve as a definitive, long-term repository for all of your sensitive data discovery results.

If you're the Macie administrator for an organization, this includes sensitive data discovery results for member accounts that you enable automated sensitive data discovery for.

To verify that you configured this repository, choose **Discovery results** in the navigation pane on the Amazon Macie console. If you prefer to do this programmatically, use the GetClassificationExportConfiguration operation of the Amazon Macie API. To learn more about sensitive data discovery results and how to configure this repository, see Storing and retaining sensitive data discovery results.

If you configured the repository, Macie creates a folder named `automated-sensitive-data-discovery` in the repository when you enable automated sensitive data discovery for the first time. This folder stores sensitive data discovery results that Macie creates while performing automated discovery for your account or organization.

If you use Macie in multiple AWS Regions, verify that you configured the repository for each of those Regions.

Step 2: Verify your permissions

To verify your permissions, use AWS Identity and Access Management (IAM) to review the IAM policies that are attached to your IAM identity. Then compare the information in those policies to the following list of actions that you must be allowed to perform:

- `macie2:GetMacieSession`
- `macie2:UpdateAutomatedDiscoveryConfiguration`
- `macie2:ListClassificationScopes`
- `macie2:UpdateClassificationScope`
- `macie2:ListSensitivityInspectionTemplates`
- `macie2:UpdateSensitivityInspectionTemplate`

The first action allows you to access your Amazon Macie account. The second action allows you to enable or disable automated sensitive data discovery for your account or organization. For an organization, it also allows you to enable automated discovery automatically for accounts in your organization. The remaining actions allow you to identify and change the configuration settings.

If you plan to review or change the configuration settings by using the Amazon Macie console, you must also be allowed to perform the following actions:

- `macie2:GetAutomatedDiscoveryConfiguration`
- `macie2:GetClassificationScope`
- `macie2:GetSensitivityInspectionTemplate`

These actions allow you to retrieve your current configuration settings and the status of automated sensitive data discovery for your account or organization. Permission to perform these actions is optional if you plan to change the configuration settings programmatically.

If you're the Macie administrator for an organization, you must also be allowed to perform the following actions:

- `macie2:ListAutomatedDiscoveryAccounts`
- `macie2:BatchUpdateAutomatedDiscoveryAccounts`

The first action allows you to retrieve the status of automated sensitive data discovery for individual accounts in your organization. The second action allows you to enable or disable automated discovery for individual accounts in your organization.

If you're not allowed to perform the requisite actions, ask your AWS administrator for assistance.

Next steps

After you complete the preceding tasks, you're ready to enable and configure the settings for your account or organization:

- Enabling automated sensitive data discovery
- Configuring settings for automated sensitive data discovery

Enabling automated sensitive data discovery

When you enable automated sensitive data discovery, Amazon Macie begins evaluating your Amazon Simple Storage Service (Amazon S3) inventory data and performing other automated discovery activities for your account in the current AWS Region. If you're the Macie administrator for an organization, by default the evaluation and activities include S3 buckets that your member accounts own. Depending on the size of your Amazon S3 data estate, statistics and other results can begin to appear within 48 hours.

After you enable automated sensitive data discovery, you can configure settings that refine the scope and nature of the analyses that Macie performs. These settings specify any S3 buckets to exclude from analyses. They also specify the managed data identifiers, custom data identifiers, and allow lists that you want Macie to use when it analyzes S3 objects. For information about these settings, see Configuring settings for automated sensitive data discovery. If you're the Macie administrator for an organization, you can also refine the scope of the analyses by enabling or disabling automated sensitive data discovery for individual accounts in your organization on a case-by-case basis.

To enable automated sensitive data discovery, you must be the Macie administrator for an organization or have a standalone Macie account. If you have a member account in an organization, work with your Macie administrator to enable automated sensitive data discovery for your account.

To enable automated sensitive data discovery

If you're the Macie administrator for an organization or you have a standalone Macie account, you can enable automated sensitive data discovery by using the Amazon Macie console or the Amazon Macie API. If you're enabling it for the first time, start by completing the prerequisite tasks. This helps ensure that you have the resources and permissions that you need.

Console

Follow these steps to enable automated sensitive data discovery by using the Amazon Macie console.

To enable automated sensitive data discovery

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to enable automated sensitive data discovery.
3. In the navigation pane, under **Settings**, choose **Automated sensitive data discovery**.
4. If you have a standalone Macie account, choose **Enable** in the **Status** section.
5. If you're the Macie administrator for an organization, choose an option in the **Status** section to specify the accounts to enable automated sensitive data discovery for:

   - To enable it for all the accounts in your organization, choose **Enable**. In the dialog box that appears, choose **My organization**. For an organization in AWS Organizations, select **Enable automatically for new accounts** to also enable it automatically for accounts that subsequently join your organization. When you finish, choose **Enable**.

- To enable it for only particular member accounts, choose **Manage accounts**. Then, in the table on the **Accounts** page, select the checkbox for each account to enable it for. When you finish, choose **Enable automated sensitive data discovery** on the **Actions** menu.
- To enable it for only your Macie administrator account, choose **Enable**. In the dialog box that appears, choose **My account** and clear **Enable automatically for new accounts**. When you finish, choose **Enable**.

If you use Macie in multiple Regions and want to enable automated sensitive data discovery in additional Regions, repeat the preceding steps in each additional Region.

To subsequently check or change the status of automated sensitive data discovery for individual accounts in an organization, choose **Accounts** in the navigation pane. On the **Accounts** page, the **Automated sensitive data discovery** field in the table indicates the current status of automated discovery for an account. To change the status for an account, select the checkbox for the account. Then use the **Actions** menu to enable or disable automated discovery for the account.

API

To enable automated sensitive data discovery programmatically, you have several options:

- To enable it for a Macie administrator account, an organization, or a standalone Macie account, use the UpdateAutomatedDiscoveryConfiguration operation. Or, if you're using the AWS Command Line Interface (AWS CLI), run the update-automated-discovery-configuration command.
- To enable it for only particular member accounts in an organization, use the BatchUpdateAutomatedDiscoveryAccounts operation. Or, if you're using the AWS CLI, run the batch-update-automated-discovery-accounts command. To enable automated discovery for a member account, you must first enable it for your administrator account or organization.

Additional options and details vary depending on the type of account that you have.

If you're a Macie administrator, use the **UpdateAutomatedDiscoveryConfiguration** operation or run the **update-automated-discovery-configuration** command to enable automated sensitive data discovery for your account or organization. In your request, specify `ENABLED` for the `status` parameter. For the `autoEnableOrganizationMembers` parameter, specify the accounts to enable it for. If you're using the AWS CLI, specify the accounts by using the `auto-enable-organization-members` parameter. Valid values are:

- `ALL` (default) – Enable it for all the accounts in your organization. This includes your administrator account, existing member accounts, and accounts that subsequently join your organization.
- `NEW` – Enable it for your administrator account. Also enable it automatically for accounts that subsequently join your organization. If you previously enabled automated discovery for your organization and you specify this value, automated discovery will continue to be enabled for existing member accounts that it's currently enabled for.
- `NONE` – Enable it for only your administrator account. Don't enable it automatically for accounts that subsequently join your organization. If you previously enabled automated discovery for your organization and you specify this value, automated discovery will continue to be enabled for existing member accounts that it's currently enabled for.

If you want to selectively enable automated sensitive data discovery for only particular member accounts, specify `NEW` or `NONE`. You can then use the **BatchUpdateAutomatedDiscoveryAccounts** operation or run the **batch-update-automated-discovery-accounts** command to enable automated discovery for the accounts.

If you have a standalone Macie account, use the **UpdateAutomatedDiscoveryConfiguration** operation or run the **update-automated-discovery-configuration** command to enable automated sensitive data discovery for your account. In your request, specify `ENABLED` for the `status` parameter. For the `autoEnableOrganizationMembers` parameter, consider whether you plan to become the Macie administrator for other accounts, and specify the appropriate value. If you specify `NONE`, automated discovery isn't enabled automatically for an account when you become the Macie administrator for the account. If you specify `ALL` or `NEW`, automated discovery is enabled automatically for the account. If you're using the AWS CLI, use the `auto-enable-organization-members` parameter to specify the appropriate value for this setting.

The following examples show how to use the AWS CLI to enable automated sensitive data discovery for one or more accounts in an organization. This first example enables automated discovery for all the accounts in an organization for the first time. It enables automated discovery for the Macie administrator account, all existing member accounts, and any accounts that subsequently join the organization.

```
$ aws macie2 update-automated-discovery-configuration --status ENABLED --auto-enable-organization-members ALL --region us-east-1
```

Where *us-east-1* is the Region in which to enable automated sensitive data discovery for the accounts, the US East (N. Virginia) Region. If the request succeeds, Macie enables automated discovery for the accounts and returns an empty response.

The next example changes the member enablement setting for an organization to NONE. With this change, automated sensitive data discovery isn't enabled automatically for accounts that subsequently join the organization. Instead, it's enabled only for the Macie administrator account, and any existing member accounts that it's currently enabled for.

```
$ aws macie2 update-automated-discovery-configuration --status ENABLED --auto-
enable-organization-members NONE --region us-east-1
```

Where *us-east-1* is the Region in which to change the setting, the US East (N. Virginia) Region. If the request succeeds, Macie updates the setting and returns an empty response.

The following examples enable automated sensitive data discovery for two member accounts in an organization. The Macie administrator has already enabled automated discovery for the organization. This example is formatted for Linux, macOS, or Unix, and it uses the backslash (\) line-continuation character to improve readability.

```
$ aws macie2 batch-update-automated-discovery-accounts \
--region us-east-1 \
--accounts '[{"accountId":"123456789012","status":"ENABLED"},
{"accountId":"111122223333","status":"ENABLED"}]'
```

This example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws macie2 batch-update-automated-discovery-accounts ^
--region us-east-1 ^
--accounts=[{\"accountId\":\"123456789012\",\"status\":\"ENABLED\"},{\"accountId\":
\"111122223333\",\"status\":\"ENABLED\"}]
```

Where:

- *us-east-1* is the Region in which to enable automated sensitive data discovery for the specified accounts, the US East (N. Virginia) Region.
- *123456789012* and *111122223333* are the account IDs for the accounts to enable automated sensitive data discovery for.

If the request succeeds for all specified accounts, Macie returns an empty `errors` array. If the request fails for some accounts, the array specifies the error that occurred for each affected account. For example:

```
"errors": [
    {
        "accountId": "123456789012",
        "errorCode": "ACCOUNT_PAUSED"
    }
]
```

In the preceding response, the request failed for the specified account (`123456789012`) because Macie is currently suspended for the account. To address this error, the Macie administrator must first enable Macie for the account.

If the request fails for all accounts, you receive a message that describes the error that occurred.

Configuring settings for automated sensitive data discovery

If you enable automated sensitive data discovery for your account or organization, you can adjust your automated discovery settings to refine the analyses that Amazon Macie performs. The settings specify Amazon Simple Storage Service (Amazon S3) buckets to exclude from analyses. They also specify the types and occurrences of sensitive data to detect and report—the managed data identifiers, custom data identifiers, and allow lists to use when analyzing S3 objects.

By default, Macie performs automated sensitive data discovery for all the S3 general purpose buckets for your account. If you're the Macie administrator for an organization, this includes buckets that your member accounts own. You can exclude specific buckets from the analyses. For example, you might exclude buckets that typically store AWS logging data, such as AWS CloudTrail event logs. If you exclude a bucket, you can include it again later.

In addition, Macie analyzes S3 objects by using only the set of managed data identifiers that we recommend for automated sensitive data discovery. Macie doesn't use custom data identifiers or allow lists that you defined. To customize the analyses, you can add or remove specific managed data identifiers, custom data identifiers, and allow lists.

If you change a setting, Macie applies your change when the next evaluation and analysis cycle starts, typically within 24 hours. In addition, your change applies only to the current AWS Region.

To make the same change in additional Regions, repeat the applicable steps in each additional Region.

Topics

> **Note**
>
> To configure settings for automated sensitive data discovery, you must be the Macie administrator for an organization or have a standalone Macie account. If your account is part of an organization, only the Macie administrator for your organization can configure and manage the settings for accounts in your organization. If you have a member account, contact your Macie administrator to learn about the settings for your account and organization.

Configuration options for organizations

If an account is part of an organization that centrally manages multiple Amazon Macie accounts, the Macie administrator for the organization configures and manages automated sensitive data discovery for accounts in the organization. This includes settings that define the scope and nature of the analyses that Macie performs for the accounts. Members can't access these settings for their own accounts.

If you're the Macie administrator for an organization, you can define the scope of the analyses in several ways:

- **Automatically enable automated sensitive data discovery for accounts** – When you enable automated sensitive data discovery, you specify whether to enable it for all existing accounts and new member accounts, only for new member accounts, or no member accounts. If you enable it for new member accounts, it's enabled automatically for any account that subsequently joins your organization, when the account joins your organization in Macie. If it's enabled for an

account, Macie includes S3 buckets that the account owns. If it's disabled for an account, Macie excludes buckets that the account owns.

- **Selectively enable automated sensitive data discovery for accounts** – With this option, you enable or disable automated sensitive data discovery for individual accounts on a case-by-case basis. If you enable it for an account, Macie includes S3 buckets that the account owns. If you don't enable it or you disable it for an account, Macie excludes buckets that the account owns.
- **Exclude specific S3 buckets from automated sensitive data discovery** – If you enable automated sensitive data discovery for an account, you can exclude particular S3 buckets that the account owns. Macie then skips the buckets when it performs automated discovery. To exclude particular buckets, add them to the exclusion list in the configuration settings for your administrator account. You can exclude as many as 1,000 buckets for your organization.

By default, automated sensitive data discovery is enabled automatically for all new and existing accounts in an organization. In addition, Macie includes all the S3 buckets that the accounts own. If you keep the default settings, this means that Macie performs automated discovery for all the buckets for your administrator account, which includes all the buckets that your member accounts own.

As a Macie administrator, you also define the nature of the analyses that Macie performs for your organization. You do this by configuring additional settings for your administrator account—the managed data identifiers, custom data identifiers, and allows lists that you want Macie to use when it analyzes S3 objects. Macie uses the settings for your administrator account when it analyzes S3 objects for other accounts in your organization.

Excluding or including S3 buckets in automated sensitive data discovery

By default, Amazon Macie performs automated sensitive data discovery for all the S3 general purpose buckets for your account. If you're the Macie administrator for an organization, this includes buckets that your member accounts own.

To refine the scope, you can exclude as many as 1,000 S3 buckets from analyses. If you exclude a bucket, Macie stops selecting and analyzing objects in the bucket when it performs automated sensitive data discovery. Existing sensitive data discovery statistics and details for the bucket persist. For example, the bucket's current sensitivity score remains unchanged. After you exclude a bucket, you can include it again later.

To exclude or include an S3 bucket in automated sensitive data discovery

You can exclude or subsequently include an S3 bucket by using the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to exclude or subsequently include an S3 bucket by using the Amazon Macie console.

To exclude or include an S3 bucket

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to exclude or include specific S3 buckets in analyses.
3. In the navigation pane, under **Settings**, choose **Automated sensitive data discovery**.

   The **Automated sensitive data discovery** page appears and displays your current settings. On that page, the **S3 buckets** section lists S3 buckets that are currently excluded, or it indicates that all buckets are currently included.
4. In the **S3 buckets** section, choose **Edit**.
5. Do one of the following:

   - To exclude one or more S3 buckets, choose **Add buckets to the exclude list**. Then, in the **S3 buckets** table, select the checkbox for each bucket to exclude. The table lists all the general purpose buckets for your account or organization in the current Region.
   - To include one or more S3 buckets that you previously excluded, choose **Remove buckets from the exclude list**. Then, in the **S3 buckets** table, select the checkbox for each bucket to include. The table lists all the buckets that are currently excluded from analyses.

   To find specific buckets more easily, enter search criteria in the search box above the table. You can also sort the table by choosing a column heading.
6. When you finish selecting buckets, choose **Add** or **Remove**, depending on the option that you chose in the preceding step.

> **Tip**
>
> You can also exclude or include individual S3 buckets on a case-by-case basis while you review bucket details on the console. To do this, choose the bucket on the **S3 buckets**

page. Then, in the details panel, change the **Exclude from automated discovery** setting for the bucket.

API

To exclude or subsequently include an S3 bucket programmatically, use the Amazon Macie API to update the classification scope for your account. The classification scope specifies buckets that you don't want Macie to analyze when it performs automated sensitive data discovery. It defines a bucket exclusion list for automated discovery.

When you update the classification scope, you specify whether to add or remove individual buckets from the exclusion list, or overwrite the current list with a new list. Therefore, it's a good idea to start by retrieving and reviewing your current list. To retrieve the list, use the GetClassificationScope operation. If you're using the AWS Command Line Interface (AWS CLI), run the get-classification-scope command to retrieve the list.

To retrieve or update the classification scope, you have to specify its unique identifier (`id`). You can get this identifier by using the GetAutomatedDiscoveryConfiguration operation. This operation retrieves your current configuration settings for automated sensitive data discovery, including the unique identifier for the classification scope for your account in the current AWS Region. If you're using the AWS CLI, run the get-automated-discovery-configuration command to retrieve this information.

When you're ready to update the classification scope, use the UpdateClassificationScope operation or, if you're using the AWS CLI, run the update-classification-scope command. In your request, use the supported parameters to exclude or include an S3 bucket in subsequent analyses:

- To exclude one or more buckets, specify the name of each bucket for the `bucketNames` parameter. For the `operation` parameter, specify `ADD`.
- To include one or more buckets that you previously excluded, specify the name of each bucket for the `bucketNames` parameter. For the `operation` parameter, specify `REMOVE`.
- To overwrite the current list with a new list of buckets to exclude, specify `REPLACE` for the `operation` parameter. For the `bucketNames` parameter, specify the name of each bucket to exclude.

Each value for the `bucketNames` parameter must be the full name of an existing general purpose bucket in the current Region. Values are case sensitive. If your request succeeds, Macie updates the classification scope and returns an empty response.

The following examples show how to use the AWS CLI to update the classification scope for an account. The first set of examples excludes two S3 buckets (*amzn-s3-demo-bucket1* and *amzn-s3-demo-bucket2*) from subsequent analyses. It adds the buckets to the list of buckets to exclude.

This example is formatted for Linux, macOS, or Unix, and it uses the backslash (\) line-continuation character to improve readability.

```
$ aws macie2 update-classification-scope \
--id 117aff7ed76b59a59c3224ebdexample \
--s3 '{"excludes":{"bucketNames":["amzn-s3-demo-bucket1","amzn-s3-demo-
bucket2"],"operation": "ADD"}}'
```

This example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws macie2 update-classification-scope ^
--id 117aff7ed76b59a59c3224ebdexample ^
--s3={\"excludes\":{\"bucketNames\":[\"amzn-s3-demo-bucket1\",\"amzn-s3-demo-
bucket2\"],\"operation\":\"ADD\"}}
```

The next set of examples later includes the buckets (*amzn-s3-demo-bucket1* and *amzn-s3-demo-bucket2*) in subsequent analyses. It removes the buckets from the list of buckets to exclude. For Linux, macOS, or Unix:

```
$ aws macie2 update-classification-scope \
--id 117aff7ed76b59a59c3224ebdexample \
--s3 '{"excludes":{"bucketNames":["amzn-s3-demo-bucket1","amzn-s3-demo-
bucket2"],"operation": "REMOVE"}}'
```

For Microsoft Windows:

```
C:\> aws macie2 update-classification-scope ^
--id 117aff7ed76b59a59c3224ebdexample ^
--s3={\"excludes\":{\"bucketNames\":[\"amzn-s3-demo-bucket1\",\"amzn-s3-demo-
bucket2\"],\"operation\":\"REMOVE\"}}
```

The following examples overwrite and replace the current list with a new list of S3 buckets to exclude. The new list specifies three buckets to exclude: *amzn-s3-demo-bucket*, *amzn-s3-demo-bucket2*, and *amzn-s3-demo-bucket3*. For Linux, macOS, or Unix:

```
$ aws macie2 update-classification-scope \
--id 117aff7ed76b59a59c3224ebdexample \
--s3 '{"excludes":{"bucketNames":["amzn-s3-demo-bucket","amzn-s3-demo-bucket2","amzn-s3-demo-bucket3"],"operation": "REPLACE"}}'
```

For Microsoft Windows:

```
C:\> aws macie2 update-classification-scope ^
--id 117aff7ed76b59a59c3224ebdexample ^
--s3={\"excludes\":{\"bucketNames\":[\"amzn-s3-demo-bucket\",\"amzn-s3-demo-bucket2\",\"amzn-s3-demo-bucket3\"],\"operation\":\"REPLACE\"}}
```

Adding or removing managed data identifiers from automated sensitive data discovery

A *managed data identifier* is a set of built-in criteria and techniques that are designed to detect a specific type of sensitive data—for example, credit card numbers, AWS secret access keys, or passport numbers for a particular country or region. By default, Amazon Macie analyzes S3 objects by using the set of managed data identifiers that we recommend for automated sensitive data discovery. To review a list of these identifiers, see Default settings for automated sensitive data discovery.

You can tailor the analyses to focus on specific types of sensitive data:

- Add managed data identifiers for the types of sensitive data that you want Macie to detect and report, and
- Remove managed data identifiers for the types of sensitive data that you don't want Macie to detect and report.

For a complete list of all the managed data identifiers that Macie currently provides and details for each one, see Using managed data identifiers.

If you remove a managed data identifier, your change doesn't affect existing sensitive data discovery statistics and details for S3 buckets. For example, if you remove the managed data identifier for AWS secret access keys and Macie previously detected that data in a bucket, Macie

continues to report those detections. However, instead of removing the identifier, which affects subsequent analyses of all buckets, consider excluding its detections from sensitivity scores for only particular buckets. For more information, see Adjusting sensitivity scores for S3 buckets.

To add or remove managed data identifiers from automated sensitive data discovery

You can add or remove managed data identifiers by using the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to add or remove a managed data identifier by using the Amazon Macie console.

To add or remove a managed data identifier

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to add or remove a managed data identifier from analyses.
3. In the navigation pane, under **Settings**, choose **Automated sensitive data discovery**.

   The **Automated sensitive data discovery** page appears and displays your current settings. On that page, the **Managed data identifiers** section displays your current settings, organized into two tabs:

   - **Added to default** – This tab lists managed data identifiers that you added. Macie uses these identifiers in addition to the ones that are in the default set and you haven't removed.
   - **Removed from default** – This tab lists managed data identifiers that you removed. Macie doesn't use these identifiers.
4. In the **Managed data identifiers** section, choose **Edit**.
5. Do any of the following:

   - To add one or more managed data identifiers, choose the **Added to default** tab. Then, in the table, select the checkbox for each managed data identifier to add. If a checkbox is already selected, you already added that identifier.
   - To remove one or more managed data identifiers, choose the **Removed from default** tab. Then, in the table, select the checkbox for each managed data identifier to remove. If a checkbox is already selected, you already removed that identifier.

On each tab, the table displays a list of all the managed data identifiers that Macie currently provides. In the table, the first column specifies each managed data identifier's ID. The ID describes the type of sensitive data that an identifier is designed to detect—for example, **USA_PASSPORT_NUMBER** for US passport numbers. To find specific managed data identifiers more easily, enter search criteria in the search box above the table. You can also sort the table by choosing a column heading.

6. When you finish, choose **Save**.

API

To add or remove a managed data identifier programmatically, use the Amazon Macie API to update the sensitivity inspection template for your account. The template stores settings that specify which managed data identifiers to use (*include*) in addition to the ones in the default set. They also specify managed data identifiers to not use (*exclude*). The settings also specify any custom data identifiers and allow lists that you want Macie to use.

When you update the template, you overwrite its current settings. Therefore, it's a good idea to start by retrieving your current settings and determining which ones you want to keep. To retrieve your current settings, use the GetSensitivityInspectionTemplate operation. If you're using the AWS Command Line Interface (AWS CLI), run the get-sensitivity-inspection-template command to retrieve the settings.

To retrieve or update the template, you have to specify its unique identifier (`id`). You can get this identifier by using the GetAutomatedDiscoveryConfiguration operation. This operation retrieves your current configuration settings for automated sensitive data discovery, including the unique identifier for the sensitivity inspection template for your account in the current AWS Region. If you're using the AWS CLI, run the get-automated-discovery-configuration command to retrieve this information.

When you're ready to update the template, use the UpdateSensitivityInspectionTemplate operation or, if you're using the AWS CLI, run the update-sensitivity-inspection-template command. In your request, use the appropriate parameters to add or remove one or more managed data identifiers from subsequent analyses:

- To start using a managed data identifier, specify its ID for the `managedDataIdentifierIds` parameter of the `includes` parameter.

- To stop using a managed data identifier, specify its ID for the `managedDataIdentifierIds` parameter of the `excludes` parameter.
- To restore the default settings, don't specify any IDs for the `includes` and `excludes` parameters. Macie then starts using only the managed data identifiers that are in the default set.

In addition to the parameters for managed data identifiers, use the appropriate `includes` parameters to specify any custom data identifiers (`customDataIdentifierIds`) and allow lists (`allowListIds`) that you want Macie to use. Also specify the Region that your request applies to. If your request succeeds, Macie updates the template and returns an empty response.

The following examples show how to use the AWS CLI to update the sensitivity inspection template for an account. The examples add one managed data identifier and remove another from subsequent analyses. They also maintain current settings that specify two custom data identifiers to use.

This example is formatted for Linux, macOS, or Unix, and it uses the backslash (\) line-continuation character to improve readability.

```
$ aws macie2 update-sensitivity-inspection-template \
--id fd7b6d71c8006fcd6391e6eedexample \
--excludes '{"managedDataIdentifierIds":["UK_ELECTORAL_ROLL_NUMBER"]}' \
--includes '{"managedDataIdentifierIds":
["STRIPE_CREDENTIALS"],"customDataIdentifierIds":
["3293a69d-4a1e-4a07-8715-208ddexample","6fad0fb5-3e82-4270-bede-469f2example"]}'
```

This example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws macie2 update-sensitivity-inspection-template ^
--id fd7b6d71c8006fcd6391e6eedexample ^
--excludes={\"managedDataIdentifierIds\":[\"UK_ELECTORAL_ROLL_NUMBER\"]} ^
--includes={\"managedDataIdentifierIds\":[\"STRIPE_CREDENTIALS\"],
\"customDataIdentifierIds\":[\"3293a69d-4a1e-4a07-8715-208ddexample\",
\"6fad0fb5-3e82-4270-bede-469f2example\"]}
```

Where:

- *fd7b6d71c8006fcd6391e6eedexample* is the unique identifier for the sensitivity inspection template to update.
- *UK_ELECTORAL_ROLL_NUMBER* is the ID for the managed data identifier to stop using (*exclude*).
- *STRIPE_CREDENTIALS* is the ID for the managed data identifier to start using (*include*).
- *3293a69d-4a1e-4a07-8715-208ddexample* and *6fad0fb5-3e82-4270-bede-469f2example* are the unique identifiers for the custom data identifiers to use.

Adding or removing custom data identifiers from automated sensitive data discovery

A *custom data identifier* is a set of criteria that you define to detect sensitive data. The criteria consist of a regular expression (*regex*) that defines a text pattern to match and, optionally, character sequences and a proximity rule that refine the results. To learn more, see Building custom data identifiers.

By default, Amazon Macie doesn't use custom data identifiers when it performs automated sensitive data discovery. If you want Macie to use specific custom data identifiers, you can add them to subsequent analyses. Macie then uses the custom data identifiers in addition to any managed data identifiers that you configure Macie to use.

If you add a custom data identifier, you can later remove it. Your change doesn't affect existing sensitive data discovery statistics and details for S3 buckets. That is to say, if you remove a custom data identifier that previously produced detections for a bucket, Macie continues to report those detections. However, instead of removing the identifier, which affects subsequent analyses of all buckets, consider excluding its detections from sensitivity scores for only particular buckets. For more information, see Adjusting sensitivity scores for S3 buckets.

To add or remove custom data identifiers from automated sensitive data discovery

You can add or remove custom data identifiers by using the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to add or remove a custom data identifier by using the Amazon Macie console.

To add or remove a custom data identifier

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to add or remove a custom data identifier from analyses.
3. In the navigation pane, under **Settings**, choose **Automated sensitive data discovery**.

   The **Automated sensitive data discovery** page appears and displays your current settings. On that page, the **Custom data identifiers** section lists custom data identifiers that you already added, or it indicates that you haven't added any custom data identifiers.
4. In the **Custom data identifiers** section, choose **Edit**.
5. Do any of the following:

   - To add one or more custom data identifiers, select the checkbox for each custom data identifier to add. If a checkbox is already selected, you already added that identifier.
   - To remove one or more custom data identifiers, clear the checkbox for each custom data identifier to remove. If a checkbox is already cleared, Macie doesn't currently use that identifier.

   > **Tip**
   >
   > To review or test the settings for a custom data identifier before you add or remove it, choose the link icon ( ) next to the identifier's name. Macie opens a page that displays the identifier's settings. To also test the identifier with sample data, enter up to 1,000 characters of text in the **Sample data** box on that page. Then choose **Test**. Macie evaluates the sample data and reports the number of matches.
6. When you finish, choose **Save**.

API

To add or remove a custom data identifier programmatically, use the Amazon Macie API to update the sensitivity inspection template for your account. The template stores settings that specify which custom data identifiers you want Macie to use when performing automated

sensitive data discovery. The settings also specify which managed data identifiers and allow lists to use.

When you update the template, you overwrite its current settings. Therefore, it's a good idea to start by retrieving your current settings and determining which ones you want to keep. To retrieve your current settings, use the GetSensitivityInspectionTemplate operation. If you're using the AWS Command Line Interface (AWS CLI), run the get-sensitivity-inspection-template command to retrieve the settings.

To retrieve or update the template, you have to specify its unique identifier (`id`). You can get this identifier by using the GetAutomatedDiscoveryConfiguration operation. This operation retrieves your current configuration settings for automated sensitive data discovery, including the unique identifier for the sensitivity inspection template for your account in the current AWS Region. If you're using the AWS CLI, run the get-automated-discovery-configuration command to retrieve this information.

When you're ready to update the template, use the UpdateSensitivityInspectionTemplate operation or, if you're using the AWS CLI, run the update-sensitivity-inspection-template command. In your request, use the `customDataIdentifierIds` parameter to add or remove one or more custom data identifiers from subsequent analyses:

- To start using a custom data identifier, specify its unique identifier for the parameter.
- To stop using a custom data identifier, omit its unique identifier from the parameter.

Use additional parameters to specify which managed data identifiers and allow lists you want Macie to use. Also specify the Region that your request applies to. If your request succeeds, Macie updates the template and returns an empty response.

The following examples show how to use the AWS CLI to update the sensitivity inspection template for an account. The examples add two custom data identifiers to subsequent analyses. They also maintain current settings that specify which managed data identifiers and allow lists to use: use the default set of managed data identifiers and one allow list.

This example is formatted for Linux, macOS, or Unix, and it uses the backslash (\) line-continuation character to improve readability.

```
$ aws macie2 update-sensitivity-inspection-template \
--id fd7b6d71c8006fcd6391e6eedexample \
```

```
--includes '{"allowListIds":["nkr81bmtu2542yyexample"],"customDataIdentifierIds":
["3293a69d-4a1e-4a07-8715-208ddexample","6fad0fb5-3e82-4270-bede-469f2example"]}'
```

This example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws macie2 update-sensitivity-inspection-template ^
--id fd7b6d71c8006fcd6391e6eedexample ^
--includes={\"allowListIds\":[\"nkr81bmtu2542yyexample\"],\"customDataIdentifierIds
\":[\"3293a69d-4a1e-4a07-8715-208ddexample\",\"6fad0fb5-3e82-4270-
bede-469f2example\"]}
```

Where:

- *fd7b6d71c8006fcd6391e6eedexample* is the unique identifier for the sensitivity inspection template to update.
- *nkr81bmtu2542yyexample* is the unique identifier for the allow list to use.
- *3293a69d-4a1e-4a07-8715-208ddexample* and *6fad0fb5-3e82-4270-bede-469f2example* are the unique identifiers for the custom data identifiers to use.

Adding or removing allow lists from automated sensitive data discovery

In Amazon Macie, an allow list defines specific text or a text pattern that you want Macie to ignore when it inspects S3 objects for sensitive data. If text matches an entry or pattern in an allow list, Macie doesn't report the text. This is the case even if the text matches the criteria of a managed or custom data identifier. To learn more, see Defining sensitive data exceptions with allow lists.

By default, Macie doesn't use allow lists when it performs automated sensitive data discovery. If you want Macie to use specific allow lists, you can add them to subsequent analyses. If you add an allow list, you can later remove it.

To add or remove allow lists from automated sensitive data discovery

You can add or remove allow lists by using the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to add or remove an allow list by using the Amazon Macie console.

To add or remove an allow list

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to add or remove an allow list from analyses.
3. In the navigation pane, under **Settings**, choose **Automated sensitive data discovery**.

   The **Automated sensitive data discovery** page appears and displays your current settings. On that page, the **Allow lists** section specifies allow lists that you already added, or it indicates that you haven't added any allow lists.
4. In the **Allow lists** section, choose **Edit**.
5. Do any of the following:

   - To add one or more allow lists, select the checkbox for each allow list to add. If a checkbox is already selected, you already added that list.
   - To remove one or more allow lists, clear the checkbox for each allow list to remove. If a checkbox is already cleared, Macie doesn't currently use that list.

   > **Tip**
   >
   > To review the settings for an allow list before you add or remove it, choose the link icon () next to the list's name. Macie opens a page that displays the list's settings. If the list specifies a regular expression (*regex*), you can also use this page to test the regex with sample data. To do this, enter up to 1,000 characters of text in the **Sample data** box, and then choose **Test**. Macie evaluates the sample data and reports the number of matches.
6. When you finish, choose **Save**.

API

To add or remove an allow list programmatically, use the Amazon Macie API to update the sensitivity inspection template for your account. The template stores settings that specify which

allow lists you want Macie to use when performing automated sensitive data discovery. The settings also specify which managed data identifiers and custom data identifiers to use.

When you update the template, you overwrite its current settings. Therefore, it's a good idea to start by retrieving your current settings and determining which ones you want to keep. To retrieve your current settings, use the GetSensitivityInspectionTemplate operation. If you're using the AWS Command Line Interface (AWS CLI), run the get-sensitivity-inspection-template command to retrieve the settings.

To retrieve or update the template, you have to specify its unique identifier (`id`). You can get this identifier by using the GetAutomatedDiscoveryConfiguration operation. This operation retrieves your current configuration settings for automated sensitive data discovery, including the unique identifier for the sensitivity inspection template for your account in the current AWS Region. If you're using the AWS CLI, run the get-automated-discovery-configuration command to retrieve this information.

When you're ready to update the template, use the UpdateSensitivityInspectionTemplate operation or, if you're using the AWS CLI, run the update-sensitivity-inspection-template command. In your request, use the `allowListIds` parameter to add or remove one or more allow lists from subsequent analyses:

- To start using an allow list, specify its unique identifier for the parameter.
- To stop using an allow list, omit its unique identifier from the parameter.

Use additional parameters to specify which managed data identifiers and custom data identifiers you want Macie to use. Also specify the Region that your request applies to. If your request succeeds, Macie updates the template and returns an empty response.

The following examples show how to use the AWS CLI to update the sensitivity inspection template for an account. The examples add an allow list to subsequent analyses. They also maintain current settings that specify which managed data identifiers and custom data identifiers to use: use the default set of managed data identifiers and two custom data identifiers.

This example is formatted for Linux, macOS, or Unix, and it uses the backslash (\) line-continuation character to improve readability.

```
$ aws macie2 update-sensitivity-inspection-template \
```

```
--id fd7b6d71c8006fcd6391e6eedexample \
--includes '{"allowListIds":["nkr81bmtu2542yyexample"],"customDataIdentifierIds":
["3293a69d-4a1e-4a07-8715-208ddexample","6fad0fb5-3e82-4270-bede-469f2example"]}'
```

This example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws macie2 update-sensitivity-inspection-template ^
--id fd7b6d71c8006fcd6391e6eedexample ^
--includes={\"allowListIds\":[\"nkr81bmtu2542yyexample\"],\"customDataIdentifierIds
\":[\"3293a69d-4a1e-4a07-8715-208ddexample\",\"6fad0fb5-3e82-4270-
bede-469f2example\"]}
```

Where:

- *fd7b6d71c8006fcd6391e6eedexample* is the unique identifier for the sensitivity inspection template to update.
- *nkr81bmtu2542yyexample* is the unique identifier for the allow list to use.
- *3293a69d-4a1e-4a07-8715-208ddexample* and *6fad0fb5-3e82-4270-bede-469f2example* are the unique identifiers for the custom data identifiers to use.

Disabling automated sensitive data discovery

You can disable automated sensitive data discovery for an account or organization at any time. If you do this, Amazon Macie stops performing all automated discovery activities for the account or organization before a subsequent evaluation and analysis cycle starts, typically within 48 hours. Additional effects vary:

- If you're a Macie administrator and you disable it for an individual account in your organization, you and the account can continue to access to all statistical data, inventory data, and other information that Macie produced and directly provided while performing automated discovery for the account. You can enable automated discovery for the account again. Macie then resumes all automated discovery activities for the account.
- If you're a Macie administrator and you disable it for your organization, you and the accounts in your organization lose access to all statistical data, inventory data, and other information that Macie produced and directly provided while performing automated discovery for your organization. For example, your S3 bucket inventory no longer includes sensitivity visualizations

or analyses statistics. You can subsequently enable automated discovery for your organization again. Macie then resumes all automated discovery activities for accounts in your organization. If you re-enable it within 30 days, you and the accounts regain access to data and information that Macie previously produced and directly provided while performing automated discovery. If you don't re-enable it within 30 days, Macie permanently deletes this data and information.

- If you disable it for your standalone Macie account, you lose access to all statistical data, inventory data, and other information that Macie produced and directly provided while performing automated discovery for your account. If you don't re-enable it within 30 days, Macie permanently deletes this data and information.

You can continue to access sensitive data findings that Macie produced while performing automated sensitive data discovery for the account or organization. Macie stores findings for 90 days. Macie also retains your configuration settings for automated discovery. In addition, data that you stored or published to other AWS services remains intact and isn't affected, such as sensitive data discovery results in Amazon S3 and finding events in Amazon EventBridge.

To disable automated sensitive data discovery

If you're the Macie administrator for an organization or you have a standalone Macie account, you can disable automated sensitive data discovery by using the Amazon Macie console or the Amazon Macie API. If you have a member account in an organization, work with your Macie administrator to disable automated discovery for your account. Only your Macie administrator can disable automated discovery for your account.

Console

Follow these steps to disable automated sensitive data discovery by using the Amazon Macie console.

To disable automated sensitive data discovery

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to disable automated sensitive data discovery.
3. In the navigation pane, under **Settings**, choose **Automated sensitive data discovery**.
4. If you're the Macie administrator for an organization, choose an option in the **Status** section to specify the accounts to disable automated sensitive data discovery for:

- To disable it for only particular member accounts, choose **Manage accounts**. Then, in the table on the **Accounts** page, select the checkbox for each account to disable it for. When you finish, choose **Disable automated sensitive data discovery** on the **Actions** menu.
- To disable it for only your Macie administrator account, choose **Disable**. In the dialog box that appears, choose **My account**, and then choose **Disable**.
- To disable it for all the accounts in your organization and your organization overall, choose **Disable**. In the dialog box that appears, choose **My organization**, and then choose **Disable**.

5. If you have a standalone Macie account, choose **Disable** in the **Status** section.

If you use Macie in multiple Regions and want to disable automated sensitive data discovery in additional Regions, repeat the preceding steps in each additional Region.

API

With the Amazon Macie API, you can disable automated sensitive data discovery in two ways. How you disable it depends partly on the type of account that you have. If you're the Macie administrator for an organization, it also depends on whether you want to disable automated discovery for only particular member accounts or your organization overall. If you disable it for your organization, you disable it for all the accounts that are currently part of your organization. If additional accounts subsequently join your organization, automated discovery is also disabled for those accounts.

To disable automated sensitive data discovery for an organization or a standalone Macie account, use the UpdateAutomatedDiscoveryConfiguration operation. Or, if you're using the AWS Command Line Interface (AWS CLI), run the update-automated-discovery-configuration command. In your request, specify `DISABLED` for the `status` parameter.

To disable automated sensitive data discovery for only particular member accounts in an organization, use the BatchUpdateAutomatedDiscoveryAccounts operation. Or, if you're using the AWS CLI, run the batch-update-automated-discovery-accounts command. In your request, use the `accountId` parameter to specify the account ID for an account that you want to disable automated discovery for. For the `status` parameter, specify `DISABLED`. To disable automated discovery for an account, Macie must currently be enabled for the account.

The following examples show how to use the AWS CLI to disable automated sensitive data discovery for one or more accounts in an organization. This first example disables automated

discovery for an organization. It disables automated discovery for the Macie administrator account and all member accounts in the organization.

```
$ aws macie2 update-automated-discovery-configuration --status DISABLED --region us-east-1
```

Where *us-east-1* is the Region in which to disable automated sensitive data discovery for the organization, the US East (N. Virginia) Region. If the request succeeds, Macie disables automated discovery for the organization and returns an empty response.

These next examples disable automated sensitive data discovery for two member accounts in an organization. This example is formatted for Linux, macOS, or Unix, and it uses the backslash (\) line-continuation character to improve readability.

```
$ aws macie2 batch-update-automated-discovery-accounts \
--region us-east-1 \
--accounts '[{"accountId":"123456789012","status":"DISABLED"},
{"accountId":"111122223333","status":"DISABLED"}]'
```

This example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws macie2 batch-update-automated-discovery-accounts ^
--region us-east-1 ^
--accounts=[{\"accountId\":\"123456789012\",\"status\":\"DISABLED\"},{\"accountId\":
\"111122223333\",\"status\":\"DISABLED\"}]
```

Where:

- *us-east-1* is the Region in which to disable automated sensitive data discovery for the specified accounts, the US East (N. Virginia) Region.
- *123456789012* and *111122223333* are the account IDs for the accounts to disable automated sensitive data discovery for.

If the request succeeds for all specified accounts, Macie returns an empty `errors` array. If the request fails for some accounts, the array specifies the error that occurred for each affected account. For example:

```
"errors": [
```

```
    {
        "accountId": "123456789012",
        "errorCode": "ACCOUNT_PAUSED"
    }
]
```

In the preceding response, the request failed for the specified account (123456789012) because Macie is currently suspended for the account.

If the request fails for all accounts, you receive a message that describes the error that occurred. For example:

```
An error occurred (ConflictException) when calling the
 BatchUpdateAutomatedDiscoveryAccounts operation: Cannot modify account states
while auto-enable is set to ALL.
```

In the preceding response, the request failed because the member enablement setting for the organization is currently configured to enable automated sensitive data discovery for all accounts (ALL). To address the error, the Macie administrator must first change this setting to NONE or NEW. For information about this setting, see Enabling automated sensitive data discovery.

Reviewing automated sensitive data discovery results

If automated sensitive data discovery is enabled, Amazon Macie automatically generates and maintains additional inventory data, statistics, and other information about the Amazon Simple Storage Service (Amazon S3) general purpose buckets for your account. If you're the Macie administrator for an organization, by default this includes S3 buckets that your member accounts own.

The additional information captures the results of automated sensitive data discovery activities that Macie has performed thus far. It also supplements other information that Macie provides about your Amazon S3 data, such as public access and encryption settings for individual S3 buckets. In addition to metadata and statistics, Macie produces records of the sensitive data it finds and the analysis that it performs—*sensitive data findings* and *sensitive data discovery results*.

As automated sensitive data discovery progresses each day, the following features and data can help you review and evaluate the results:

- **Summary dashboard** – Provides aggregated statistics for your Amazon S3 data estate. The statistics include data for key metrics such as the total number of buckets that Macie has found sensitive data in, and how many of those buckets are publicly accessible. They also report issues that affect coverage of your Amazon S3 data.
- **S3 buckets heat map** – Provides an interactive, visual representation of data sensitivity across your data estate, grouped by AWS account. For each account, the map includes aggregated sensitivity statistics and it uses colors to indicate the current sensitivity score for each bucket that the account owns. The map also uses symbols to help you identify buckets that are publicly accessible, can't be analyzed by Macie, and more.
- **S3 buckets table** – Provides summary information for each S3 bucket in your inventory. For each bucket, the table includes data such as the bucket's current sensitivity score, the number of objects that Macie can analyze in the bucket, and whether you configured any sensitive data discovery jobs to periodically analyze objects in the bucket. You can export data from the table to a comma-separated values (CSV) file.
- **S3 bucket details** – Provides detailed statistics and information about an S3 bucket. The details include a list of objects that Macie has analyzed in the bucket, and a breakdown of the types and number of occurrences of sensitive data that Macie has found in the bucket. These are in addition to details about settings that affect the security and privacy of the bucket's data.
- **Sensitive data findings** – Provide detailed reports of sensitive data that Macie found in individual S3 objects. The details include when Macie found the sensitive data, and the types and number of occurrences of the sensitive data that Macie found. The details also include information about the affected S3 bucket and object, including the bucket's public access settings and when the object was most recently changed.
- **Sensitive data discovery results** – Provide records of the analysis that Macie performed for individual S3 objects. This includes objects that Macie doesn't find sensitive data in, and objects that Macie can't analyze due to issues or errors. If Macie finds sensitive data in an object, the sensitive data discovery result provides information about the sensitive data that Macie found.

With this data, you can evaluate data sensitivity across your Amazon S3 data estate and drill down to evaluate and investigate individual S3 buckets and objects. Combined with information that Macie provides about the security and privacy of your Amazon S3 data, you can also identify cases where immediate remediation might be necessary—for example, a publicly accessible bucket that Macie found sensitive data in.

Additional data can help you assess and monitor coverage of your Amazon S3 data. With coverage data, you can check the status of the analyses for your data estate overall and individual S3

buckets within it. You can also identify issues that prevented Macie from analyzing objects in specific buckets. If you remediate the issues, you can increase coverage of your Amazon S3 data during subsequent analysis cycles. For more information, see Assessing automated sensitive data discovery coverage.

Topics

Reviewing data sensitivity statistics on the Summary dashboard

On the Amazon Macie console, the **Summary** dashboard provides a snapshot of aggregated statistics and findings data for your Amazon Simple Storage Service (Amazon S3) data in the current AWS Region. It's designed to help you assess the overall security posture of your Amazon S3 data.

Dashboard statistics include data for key security metrics such as the number of S3 general purpose buckets that are publicly accessible or shared with other AWS accounts. The dashboard also displays groups of aggregated findings data for your account—for example, the buckets that generated the most findings during the preceding seven days. If you're the Macie administrator for an organization, the dashboard provides aggregated statistics and data for all the accounts in your organization. You can optionally filter the data by account.

If automated sensitive data discovery is enabled, the **Summary** dashboard includes additional statistics. The statistics capture the status and results of automated discovery activities that Macie has performed thus far for your Amazon S3 data. The following image shows an example of these statistics.

The statistics are organized primarily into two sections, **Automated discovery** and **Coverage issues**. Statistics in the **Automated discovery** section provide a snapshot of the current status and results of automated sensitive data discovery activities. Statistics in the **Coverage issues** section indicate whether issues prevented Macie from analyzing objects in individual S3 buckets. The statistics don't include data for sensitive data discovery jobs that you create and run. However, remediating coverage issues for automated sensitive data discovery is likely to also increase coverage by jobs that you subsequently run.

Topics

- Displaying the Summary dashboard
- Understanding sensitive data discovery statistics on the Summary dashboard

Displaying the Summary dashboard

Follow these steps to display the **Summary** dashboard on the Amazon Macie console. To query the statistics programmatically, use the GetBucketStatistics operation of the Amazon Macie API.

To display the Summary dashboard

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Summary**. Macie displays the **Summary** dashboard.
3. To drill down and review the supporting data for an item on the dashboard, choose the item.

If you're the Macie administrator for an organization, the dashboard displays aggregated statistics and data for your account and member accounts in your organization. To display data for only a particular account, enter the account's ID in the **Account** box above the dashboard.

Understanding sensitive data discovery statistics on the Summary dashboard

The **Summary** dashboard includes aggregated statistics that can help you monitor automated sensitive data discovery for your Amazon S3 data. It provides a snapshot of the current status and results of the analyses for your Amazon S3 data in the current AWS Region. For example, you can use dashboard statistics to quickly determine how many S3 buckets Amazon Macie has found sensitive data in, and how many of those buckets are publicly accessible. You can also assess coverage of your Amazon S3 data. Coverage statistics can help you identify issues that prevent Macie from analyzing objects in individual S3 buckets.

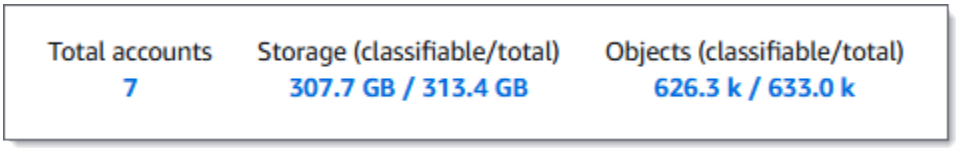
On the dashboard, statistics for automated sensitive data discovery are organized into the following sections:

- Storage and sensitive data discovery
- Automated discovery
- Coverage issues

Individual statistics in each section are as follows. For information about statistics in other sections of the dashboard, see Understanding components of the Summary dashboard.

Storage and sensitive data discovery

At the top of the dashboard, statistics indicate how much data you store in Amazon S3, and how much of that data Amazon Macie can analyze to detect sensitive data. The following image shows an example of these statistics for an organization with seven accounts.



Individual statistics in this section are:

- **Total accounts** – This field appears if you're the Macie administrator for an organization or you have a standalone Macie account. It indicates the total number of AWS accounts that own buckets in your bucket inventory. If you're a Macie administrator, this is the total number of Macie accounts that you manage for your organization. If you have a standalone Macie account, this value is *1*.

Total S3 buckets – This field appears if you have a member account in an organization. It indicates the total number of general purpose buckets in your inventory, including buckets that don't store any objects.

- **Storage** – These statistics provide information about the storage size of objects in your bucket inventory:
  - **Classifiable** – The total storage size of all the objects that Macie can analyze in the buckets.
  - **Total** – The total storage size of all the objects in the buckets, including objects that Macie can't analyze.

  If any of the objects are compressed files, these values don't reflect the actual size of those files after they're decompressed. If versioning is enabled for any of the buckets, these values are based on the storage size of the latest version of each object in those buckets.
- **Objects** – These statistics provide information about the number of objects in your bucket inventory:
  - **Classifiable** – The total number of objects that Macie can analyze in the buckets.
  - **Total** – The total number of objects in the buckets, including objects that Macie can't analyze.

In the preceding statistics, data and objects are *classifiable* if they use a supported Amazon S3 storage class and they have a file name extension for a supported file or storage format. You can detect sensitive data in the objects by using Macie. For more information, see Supported storage classes and formats.

Note that **Storage** and **Objects** statistics don't include data about objects in buckets that Macie isn't allowed to access. To identify buckets where this is the case, choose the **Access denied** statistic in the **Coverage issues** section of the dashboard.

Automated discovery

This section captures the status and results of automated sensitive data discovery activities that Amazon Macie has performed thus far for your Amazon S3 data. The following image shows an example of the statistics that this section provides.

Individual statistics in this section are as follows.

Total buckets

The doughnut chart indicates the total number of buckets in your bucket inventory. The chart groups the buckets into categories based on each bucket's current sensitivity score:

- **Sensitive** (*red*) – The total number of buckets whose sensitivity score ranges from *51* through *100*.
- **Not sensitive** (*blue*) – The total number of buckets whose sensitivity score ranges from *1* through *49*.
- **Not yet analyzed** (*light gray*) – The total number of buckets whose sensitivity score is *50*.
- **Classification error** (*dark gray*) – The total number of buckets whose sensitivity score is *-1*.

For details about the range of sensitivity scores and labels that Macie defines, see Sensitivity scoring for S3 buckets.

To review additional statistics for a group, hover over the group:

- **Buckets** – The total number of buckets.
- **Publicly accessible** – The total number of buckets that allow the general public to have read or write access to the bucket.
- **Classifiable bytes** – The total storage size of all the objects that Macie can analyze in the buckets. These objects use supported Amazon S3 storage classes and they have file name

extensions for supported file or storage formats. For more information, see Supported storage classes and formats.

- **Total bytes** – The total storage size of all the buckets.

In the preceding statistics, storage size values are based on the storage size of the latest version of each object in the buckets. If any of the objects are compressed files, these values don't reflect the actual size of those files after they're decompressed.

Sensitive

This area indicates the total number of buckets that currently have a sensitivity score ranging from *51* through *100*. Within this group, **Publicly accessible** indicates the total number of buckets that also allow the general public to have read or write access to the bucket.

Not sensitive

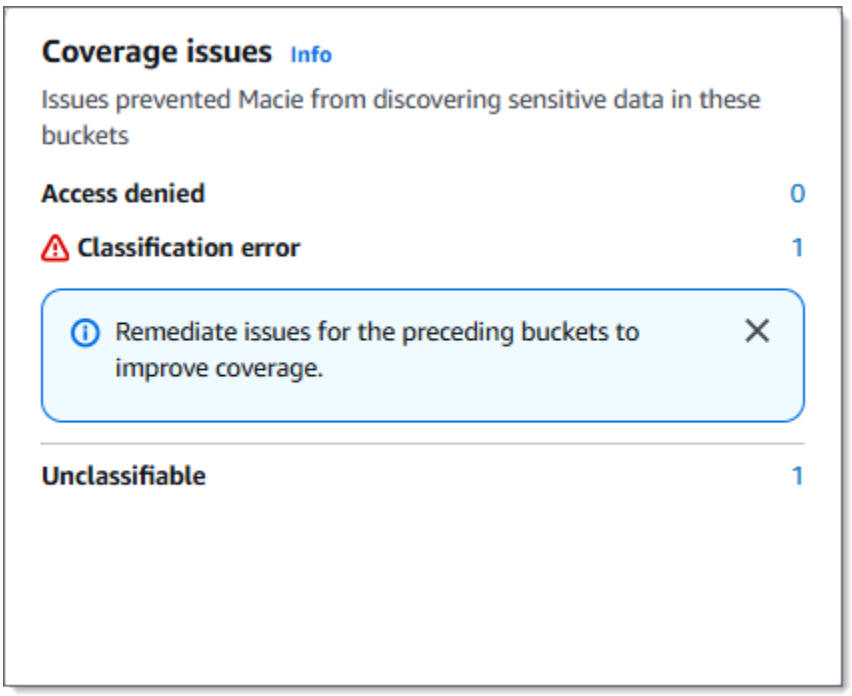
This area indicates the total number of buckets that currently have a sensitivity score ranging from *1* through *49*. Within this group, **Publicly accessible** indicates the total number of buckets that also allow the general public to have read or write access to the bucket.

To determine and calculate values for **Publicly accessible** statistics, Macie analyzes a combination of account- and bucket-level settings for each bucket, such as the block public access settings for the account and bucket, and the bucket policy for the bucket. Macie does this for up to 10,000 buckets for an account. For more information, see How Macie monitors Amazon S3 data security.

Note that statistics in the **Automated discovery** section don't include the results of sensitive data discovery jobs that you create and run.

Coverage issues

In this section, statistics indicate whether certain types of issues prevented Amazon Macie from analyzing objects in individual S3 buckets. The following image shows an example of the statistics that this section provides.



Individual statistics in this section are:

- **Access denied** – The total number of buckets that Macie isn't allowed to access. Macie can't analyze any objects in these buckets. The buckets' permissions settings prevent Macie from accessing the buckets and the buckets' objects.
- **Classification error** – The total number of buckets that Macie hasn't analyzed yet due to object-level classification errors. Macie tried to analyze one or more objects in these buckets. However, Macie couldn't analyze the objects due to issues with object-level permissions settings, object content, or quotas.
- **Unclassifiable** – The total number of buckets that don't store any classifiable objects. Macie can't analyze any objects in these buckets. All the objects use Amazon S3 storage classes that Macie doesn't support, or they have file name extensions for file or storage formats that Macie doesn't support.

Choose the value for a statistic to display additional details and, as applicable, remediation guidance. If you remediate access issues and classification errors, you can increase coverage of your Amazon S3 data during subsequent analysis cycles. For more information, see Assessing automated sensitive data discovery coverage.

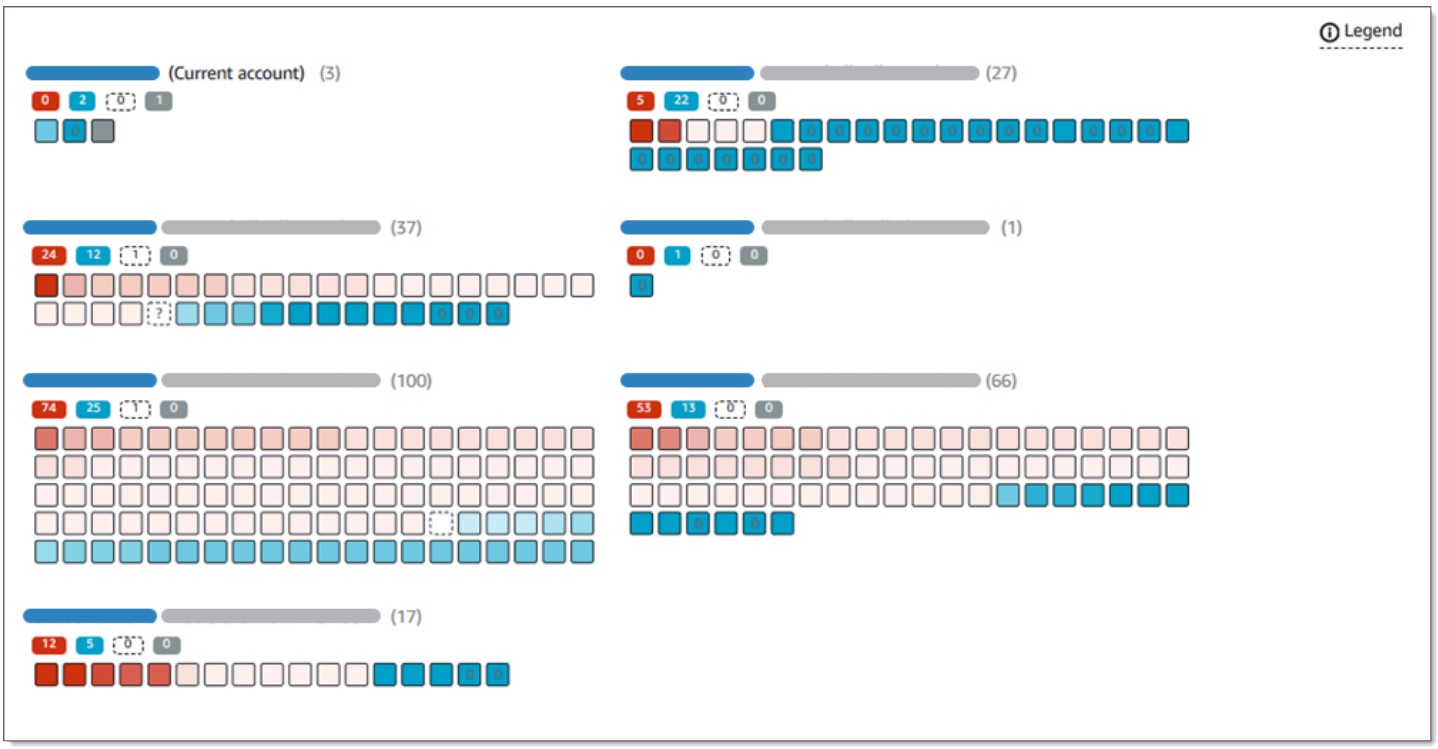
Note that statistics in the **Coverage issues** section don't explicitly include data for sensitive data discovery jobs that you create and run. However, remediating coverage issues that affect

automated sensitive data discovery is likely to also increase coverage by jobs that you subsequently run.

Visualizing data sensitivity with the S3 buckets map

On the Amazon Macie console, the **S3 buckets** heat map provides an interactive, visual representation of data sensitivity across your Amazon Simple Storage Service (Amazon S3) data estate. It captures the results of automated sensitive data discovery activities that Macie has performed thus far for your Amazon S3 data in the current AWS Region.

If you're the Macie administrator for an organization, the map includes results for S3 buckets that your member accounts own. The data is grouped by AWS account and sorted by account ID, as shown in the following image.



The map displays data for up to 100 S3 buckets for each account. To display data for all buckets, you can switch to table view and review the data in tabular format instead.

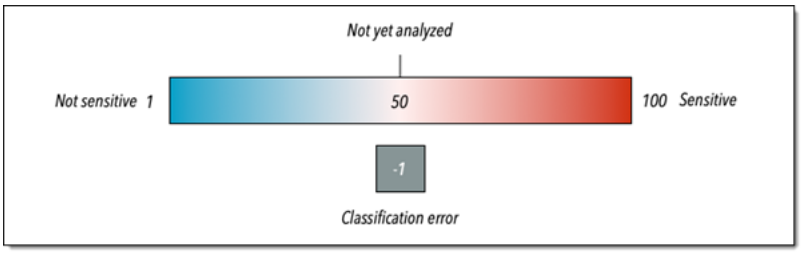
To display the map, choose **S3 buckets** in the navigation pane on the console. Then choose map ( ) at the top of the page. The map is available only if automated sensitive data discovery is currently enabled. It doesn't include the results of sensitive data discovery jobs that you create and run.

Topics

- Interpreting data in the S3 buckets map
- Interacting with the S3 buckets map

Interpreting data in the S3 buckets map

In the **S3 buckets** map, each square represents an S3 general purpose bucket in your bucket inventory. The color of a square represents a bucket's current sensitivity score, which measures the intersection of two primary dimensions: the amount of sensitive data that Macie has found in the bucket, and the amount of data that Macie has analyzed in the bucket. The intensity of the color's hue represents where a score falls in a range of data sensitivity values, as shown in the following image.



In general, you can interpret color and hue intensity as follows:

- **Blue** – If a bucket's current sensitivity score ranges from *1* through *49*, the bucket's square is blue and the bucket's sensitivity label is **Not sensitive**. The intensity of the blue hue reflects the number of unique objects that Macie has analyzed in the bucket relative to the total number of unique objects in the bucket. A darker hue indicates a lower sensitivity score.
- **No color** – If a bucket's current sensitivity score is *50*, the bucket's square isn't colored and the bucket's sensitivity label is **Not yet analyzed**. In addition, the square has a dashed border.
- **Red** – If a bucket's current sensitivity score ranges from *51* through *100*, the bucket's square is red and the bucket's sensitivity label is **Sensitive**. The intensity of the red hue reflects the amount of sensitive data that Macie has found in the bucket. A darker hue indicates a higher sensitivity score.
- **Gray** – If a bucket's current sensitivity score is *-1*, the bucket's square is dark gray and the bucket's sensitivity label is **Classification error**. Hue intensity doesn't vary.

For details about the range of sensitivity scores and labels that Macie defines, see Sensitivity scoring for S3 buckets.

In the map, the square for an S3 bucket might also contain a symbol. The symbol indicates an error, issue, or other type of consideration that might affect your evaluation of a bucket's sensitivity. A symbol can also indicate a potential issue with the security of the bucket—for example, the bucket is publicly accessible. The following table lists the symbols that Macie uses to notify you of these cases.

Symbol	Definition	Description
	**Access denied**	Macie isn't allowed to access the bucket or the bucket's objects. Consequently, Macie can't analyze any objects in the bucket. This issue typically occurs because a bucket has a restrictive bucket policy. For information about how to address this issue, see Allowing Macie to access S3 buckets and objects.
	**Publicly accessible**	The general public has read or write access to the bucket. To make this determination, Macie analyzes a combinati on of settings for each bucket, such as the block public access settings for the account and the bucket, and the bucket policy for the bucket. Macie can do this for up to 10,000 buckets for an account. For more informati on, see How Macie monitors Amazon S3 data security.

Symbol	Definition	Description
?	**Unclassifiable**	Macie can't analyze any objects in the bucket. All the bucket's objects use Amazon S3 storage classes that Macie doesn't support, or they have file name extensions for file or storage formats that Macie doesn't support. For Macie to analyze an object, the object must use a supported storage class and have a file name extension for a supported file or storage format. For more information, see Supported storage classes and formats.
0	**Zero bytes**	The bucket doesn't store any objects for Macie to analyze. The bucket is empty or all the objects in the bucket contain zero (0) bytes of data.

Interacting with the S3 buckets map

As you review the **S3 buckets** map, you can interact with it in different ways to reveal and evaluate additional data and details for individual accounts and buckets. Follow these steps to display the map and use various features that it provides.

To interact with the S3 buckets map

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **S3 buckets**. The **S3 buckets** page displays a map of your bucket inventory. If the page displays your inventory in tabular format instead, choose map

()
at the top of the page.

By default, the map doesn't display data for buckets that are currently excluded from automated sensitive data discovery. If you're the Macie administrator for an organization, it also doesn't display data for accounts that automated sensitive data discovery is currently disabled for. To display this data, choose **X** in the **Is monitored by automated discovery** filter token below the filter box.

3. At the top of the page, optionally choose refresh
()
to retrieve the latest bucket metadata from Amazon S3.
4. In the **S3 buckets** map, do any of the following:

   - To determine how many buckets have a specific sensitivity label, refer to the colored badges immediately below an AWS account ID. The badges display aggregated bucket counts, broken down by sensitivity label.

     For example, the red badge reports the total number of buckets that are owned by the account and have the **Sensitive** label. The sensitivity score for these buckets ranges from *51* through *100*. The blue badge reports the total number of buckets that are owned by the account and have the **Not sensitive** label. The sensitivity score for these buckets ranges from *1* through *49*.
   - To review a subset of information about a bucket, hover over the bucket's square. A popover displays the bucket's name and current sensitivity score.

     The popover also displays the total number of objects that Macie can analyze in the bucket and the total storage size of the latest version of those objects. These objects are *classifiable*. They use supported Amazon S3 storage classes and they have file name extensions for supported file or storage formats. For more information, see Supported storage classes and formats.
   - To filter the map and display only those buckets that have a specific value for a field, place your cursor in the filter box, and then add a filter condition for the field. Macie applies the condition's criteria and displays the condition below the filter box. To further refine the results, add filter conditions for additional fields. For more information, see Filtering your S3 bucket inventory.

- To drill down and display only those buckets that are owned by a particular account, choose the account ID for the account. Macie opens a new tab that filters and displays data only for that account.

5. To review data sensitivity statistics and other information for a particular bucket, choose the bucket's square. Then refer to the details panel. For information about these details, see Reviewing data sensitivity details for S3 buckets.

> **Tip**
>
> On the **Bucket details** tab of the panel, you can pivot and drill down on many of the fields. To show buckets that have the same value for a field, choose  in the field. To show buckets that have other values for a field, choose  in the field.

Assessing data sensitivity with the S3 buckets table

To review summary information for your Amazon Simple Storage Service (Amazon S3) buckets, you can use the **S3 buckets** table on the Amazon Macie console. By using the table, you can review and analyze an inventory of your general purpose buckets in the current AWS Region, and drill down to review detailed information and statistics for individual buckets. If you're the Macie administrator for an organization, the table includes information about buckets that your member accounts own. If you prefer to access and query the data programmatically, you can use the DescribeBuckets operation of the Amazon Macie API.

On the console, you can sort and filter the table to customize your view. You can also export data from the table to a comma-separated values (CSV) file. If you choose an S3 bucket in the table, the details panel displays additional information about the bucket. This includes details and statistics for settings and metrics that provide insight into the security and privacy of the bucket's data. If automated sensitive data discovery is enabled, it also includes data that captures the results of automated discovery activities that Macie has performed thus far for the bucket.

To assess data sensitivity by using the S3 buckets table

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.

2. In the navigation pane, choose **S3 buckets**. The **S3 buckets** page displays your bucket inventory.

   By default, the page doesn't display data for buckets that are currently excluded from automated sensitive data discovery. If you're the Macie administrator for an organization, it also doesn't display data for accounts that automated sensitive data discovery is currently disabled for. To display this data, choose **X** in the **Is monitored by automated discovery** filter token below the filter box.

3. Choose table
   ()
   at the top of the page. Macie displays the number of buckets in your inventory and a table of the buckets.

4. To retrieve the latest bucket metadata from Amazon S3, choose refresh
   ()
   at the top of the page.

   If the information icon
   ()
   appears next to any bucket names, we recommend that you do this. This icon indicates that a bucket was created during the past 24 hours, possibly after Macie last retrieved bucket and object metadata from Amazon S3 as part of the daily refresh cycle.

5. In the **S3 buckets** table, review summary information about each bucket in your inventory:

   - **Sensitivity** – The bucket's current sensitivity score. For information about the range of sensitivity scores that Macie defines, see Sensitivity scoring for S3 buckets.
   - **Bucket** – The name of the bucket.
   - **Account** – The account ID for the AWS account that owns the bucket.
   - **Classifiable objects** – The total number of objects that Macie can analyze to detect sensitive data in the bucket.
   - **Classifiable size** – The total storage size of all the objects that Macie can analyze to detect sensitive data in the bucket.

     This value doesn't reflect the actual size of any compressed objects after they're decompressed. Also, if versioning is enabled for the bucket, this value is based on the storage size of the latest version of each object in the bucket.

- **Monitored by job** – Whether you configured any sensitive data discovery jobs to periodically analyze objects in the bucket on a daily, weekly, or monthly basis.

  If the value for this field is *Yes*, the bucket is explicitly included in a periodic job or the bucket matched the criteria for a periodic job within the past 24 hours. In addition, the status of at least one of those jobs is not *Cancelled*. Macie updates this data on a daily basis.

- **Latest job run** – If you configured any one-time or periodic sensitive data discovery jobs to analyze objects in the bucket, this field indicates the most recent date and time when one of those jobs started to run. Otherwise, a dash (–) appears in this field.

In the preceding data, objects are *classifiable* if they use a supported Amazon S3 storage class and they have a file name extension for a supported file or storage format. You can detect sensitive data in the objects by using Macie. For more information, see Supported storage classes and formats.

6. To analyze your inventory by using the table, do any of the following:

- To sort the table by a specific field, choose the column heading for the field. To change the sort order, choose the column heading again.
- To filter the table and display only those buckets that have a specific value for a field, place your cursor in the filter box, and then add a filter condition for the field. To further refine the results, add filter conditions for additional fields. For more information, see Filtering your S3 bucket inventory.
- To review data sensitivity statistics and other information for a particular bucket, choose the bucket's name. Then refer to the details panel. For information about these details, see Reviewing S3 bucket details.

> **Tip**
>
> On the **Bucket details** tab of the panel, you can pivot and drill down on many of the fields. To show buckets that have the same value for a field, choose
> 
> in the field. To show buckets that have other values for a field, choose
> 
> in the field.

7. To export data from the table to a CSV file, select the checkbox for each row to export, or select the checkbox in the selection column heading to select all rows. Then choose **Export to CSV** at the top of the page. You can export up to 50,000 rows from the table.
8. To perform deeper, more immediate analysis of objects in one or more buckets, select the checkbox for each bucket. Then choose **Create job**. For more information, see Creating a sensitive data discovery job.

Reviewing data sensitivity details for S3 buckets

As automated sensitive data discovery progresses, you can review detailed results in statistics and other information that Amazon Macie provides about each of your Amazon Simple Storage Service (Amazon S3) buckets. If you're the Macie administrator for an organization, this includes buckets that your member accounts own.

The statistics and information include details that provide insight into the security and privacy of an S3 bucket's data. They also capture the results of automated sensitive data discovery activities that Macie has performed thus far for a bucket. For example, you can find a list of objects that Macie has analyzed in a bucket. You can also find a breakdown of the types and number of occurrences of sensitive data that Macie has found in a bucket. Note that this data doesn't include the results of sensitive data discovery jobs that you create and run.

Macie automatically recalculates and updates statistics and details for your S3 buckets while it performs automated sensitive data discovery. For example:

- If Macie doesn't find sensitive data in an S3 object, Macie decreases the bucket's sensitivity score and updates the bucket's sensitivity label as necessary. Macie also adds the object to the list of objects that it selected for analysis.
- If Macie finds sensitive data in an S3 object, Macie adds those occurrences to the breakdown of sensitive data types that Macie has found in the bucket. Macie also increases the bucket's sensitivity score and updates the bucket's sensitivity label as necessary. In addition, Macie adds the object to the list of objects that it selected for analysis. These tasks are in addition to creating a sensitive data finding for the object.
- If Macie finds sensitive data in an S3 object that's subsequently changed or deleted, Macie removes sensitive data occurrences for the object from the bucket's breakdown of sensitive data types. Macie also decreases the bucket's sensitivity score and updates the bucket's sensitivity label as necessary. In addition, Macie removes the object from the list of objects that it selected for analysis.

- If Macie attempts to analyze an S3 object but an issue or error prevents analysis, Macie adds the object to the list of objects that it selected for analysis, and indicates that it wasn't able to analyze the object.

If you're the Macie administrator for an organization or you have a standalone Macie account, you can optionally use these details to assess and adjust certain automated discovery settings for an S3 bucket. For example, you can include or exclude specific types of sensitive data from a bucket's score. For more information, see Adjusting sensitivity scores for S3 buckets.

To review data sensitivity details for an S3 bucket

To review data sensitivity and other details for an S3 bucket, you can use the Amazon Macie console or the Amazon Macie API. On the console, the details panel provides centralized access to this information. With the API, you can retrieve and process the data programmatically.

Console

Follow these steps to review data sensitivity and other details for an S3 bucket by using the Amazon Macie console.

To review the details for an S3 bucket

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **S3 buckets**. The **S3 buckets** page displays an interactive map of your bucket inventory. Optionally choose table () at the top of the page to display your inventory in tabular format instead.

   By default, the page doesn't display data for buckets that are currently excluded from automated sensitive data discovery. If you're the Macie administrator for an organization, it also doesn't display data for accounts that automated sensitive data discovery is currently disabled for. To display this data, choose **X** in the **Is monitored by automated discovery** filter token below the filter box.
3. To retrieve the latest bucket metadata from Amazon S3, choose refresh () at the top of the page.
4. Choose the bucket whose details you want to review. The details panel displays data sensitivity statistics and other information about the bucket.

The top of the panel shows general information about the bucket: the bucket's name, the account ID for the AWS account that owns the bucket, and the bucket's current sensitivity score. If you're a Macie administrator or you have a standalone Macie account, it also provides options for changing certain automated discovery settings for the bucket. Additional settings and information are organized into the following tabs:

Sensitivity | Bucket details | Object samples | Sensitive data discovery

Individual settings and information on each tab are as follows.

Sensitivity

This tab shows the bucket's current sensitivity score, ranging from *-1* to *100*. For information about the range of sensitivity scores that Macie defines, see Sensitivity scoring for S3 buckets.

The tab also provides a breakdown of the types of sensitive data that Macie has found in the bucket's objects, and the number of occurrences of each type:

- **Sensitive data type** – The unique identifier (ID) for the managed data identifier that detected the data, or the name of the custom data identifier that detected the data.

  A managed data identifier's ID describes the type of sensitive data that it's designed to detect—for example, **USA_PASSPORT_NUMBER** for US passport numbers. For details about each managed data identifier, see Using managed data identifiers.
- **Count** – The total number of occurrences of the data that the managed or custom data identifier detected.
- **Scoring status** – This field appears if you're a Macie administrator or you have a standalone Macie account. It specifies whether occurrences of the data are included or excluded from the bucket's sensitivity score.

  If Macie calculates the bucket's score, you can adjust the calculation by including or excluding specific types of sensitive data from the score: select the checkbox for the identifier that detected the sensitive data to include or exclude, and then choose an option on the **Actions** menu. For more information, see Adjusting sensitivity scores for S3 buckets.

If Macie hasn't found sensitive data in objects that the bucket currently stores, this section shows the **No detections found** message.

Note that the **Sensitivity** tab doesn't include data for objects that were changed or deleted after Macie analyzed them. If objects are changed or deleted after analysis, Macie automatically recalculates and updates the appropriate statistics and data to exclude the objects.

Bucket details

This tab provides details about the bucket's settings, including data security and privacy settings. For example, you can review breakdowns of the bucket's public access settings, and determine whether the bucket replicates objects or is shared with other AWS accounts.

Of special note, the **Last updated** field indicates when Macie most recently retrieved metadata from Amazon S3 for the bucket or the bucket's objects. The **Latest automated discovery run** field indicates when Macie most recently analyzed objects in the bucket while performing automated sensitive data discovery. If this analysis hasn't occurred, a dash (–) appears in this field.

The tab also provides object-level statistics that can help you assess how much data Macie can analyze in the bucket. It also indicates whether you configured any sensitive data discovery jobs to analyze objects in the bucket. If you have, you can access details about the job that ran most recently and then optionally display any findings that the job produced.

In certain cases, this tab might not include all the details of a bucket. This can occur if you store more than 10,000 buckets in Amazon S3. Macie maintains complete inventory data for only 10,000 buckets for an account—the 10,000 buckets that were most recently created or changed. Macie can, however, analyze objects in buckets that exceed this quota. To review additional details for the buckets, use Amazon S3.

For additional details about the information on this tab, see Reviewing the details of S3 buckets.

Object samples

This tab lists objects that Macie selected for analysis while performing automated sensitive data discovery for the bucket. Optionally choose an object's name to open the Amazon S3 console and display the object's properties.

The list includes data for up to 100 objects. The list is populated based on the value for the **Object sensitivity** field: **Sensitive**, followed by **Not Sensitive**, followed by objects that Macie wasn't able to analyze.

In the list, the **Object sensitivity** field indicates whether Macie found sensitive data in an object:

- **Sensitive** – Macie found at least one occurrence of sensitive data in the object.
- **Not sensitive** – Macie didn't find sensitive data in the object.
- **–** (*dash*) – Macie wasn't able to complete its analysis of the object due to an issue or error.

The **Classification result** field indicates whether Macie was able to analyze an object:

- **Complete** – Macie completed its analysis of the object.
- **Partial** – Macie analyzed only a subset of data in the object due to an issue or error. For example, the object is an archive file that contains files in an unsupported format.
- **Skipped** – Macie wasn't able to analyze any data in the object due to an issue or error. For example, the object is encrypted with a key that Macie isn't allowed to use.

Note that the list doesn't include objects that were changed or deleted after Macie analyzed or attempted to analyze them. Macie automatically removes an object from the list if the object is subsequently changed or deleted.

Sensitive data discovery

This tab provides aggregated, automated sensitive data discovery statistics for the bucket:

- **Analyzed bytes** – The total amount of data, in bytes, that Macie has analyzed in the bucket.
- **Classifiable bytes** – The total storage size, in bytes, of all the objects that Macie can analyze in the bucket. These objects use supported Amazon S3 storage classes and they have file name extensions for supported file or storage formats. For more information, see Supported storage classes and formats.
- **Total detections** – The total number of occurrences of sensitive data that Macie has found in the bucket. This includes occurrences that are currently suppressed by the sensitivity scoring settings for the bucket.

The **Objects analyzed** chart indicates the total number of objects that Macie has analyzed in the bucket. It also provides a visual representation of the number of objects that Macie did or didn't find sensitive data in. The legend below the chart shows a breakdown of these results:

- **Sensitive objects** (*red*) – The total number of objects that Macie found at least one occurrence of sensitive data in.

- **Not sensitive objects** (*blue*) – The total number of objects that Macie didn't find sensitive data in.
- **Objects skipped** (*dark gray*) – The total number of objects that Macie wasn't able to analyze due to an issue or error.

The area below the chart's legend provides a breakdown of cases where Macie wasn't able to analyze objects because certain types of permissions issues or cryptographic errors occurred:

- **Skipped: Invalid encryption** – The total number of objects that are encrypted with customer-provided keys. Macie can't access these keys.
- **Skipped: Invalid KMS** – The total number of objects that are encrypted with AWS Key Management Service (AWS KMS) keys that are no longer available. These objects are encrypted with AWS KMS keys that were disabled, are scheduled for deletion, or were deleted. Macie can't use these keys.
- **Skipped: Permission denied** – The total number of objects that Macie isn't allowed to access due to the permissions settings for the object, or the permissions settings for the key that was used to encrypt the object.

For details about these and other types of issues and errors that can occur, see Remediating coverage issues. If you remediate the issues and errors, you can increase coverage of the bucket's data during subsequent analysis cycles.

Statistics on the **Sensitive data discovery** tab don't include data for objects that were changed or deleted after Macie analyzed or attempted to analyze them. If objects are changed or deleted after Macie analyzes or attempts to analyze them, Macie automatically recalculates these statistics to exclude the objects.

API

To retrieve data sensitivity and other details for an S3 bucket programmatically, you have several options. The appropriate option depends on the details that you want to retrieve:

- To retrieve a bucket's current sensitivity score and aggregated analysis statistics, use the GetResourceProfile operation. Or, if you're using the AWS Command Line Interface (AWS CLI), run the get-resource-profile command. The statistics include data such as the number of objects that Macie has analyzed, and the number of objects that Macie has found sensitive data in.

- To retrieve a breakdown of the types and amount of sensitive data that Macie has found in a bucket, use the ListResourceProfileDetections operation. Or, if you're using the AWS CLI, run the list-resource-profile-detections command. The breakdown also provides details about the managed or custom data identifier that detected each type of sensitive data.
- To retrieve a list of up to 100 objects that Macie selected from a bucket for analysis, use the ListResourceProfileArtifacts operation. Or, if you're using the AWS CLI, run the list-resource-profile-artifacts command. For each object, the list specifies: the Amazon Resource Name (ARN) of the object, whether Macie completed its analysis of the object; and, whether Macie found sensitive data in the object.

In your request, use the `resourceArn` parameter to specify the ARN of the bucket to retrieve the details for. If you're using the AWS CLI, use the `resource-arn` parameter to specify the ARN.

For additional details about an S3 bucket, such as the bucket's public access settings, use the DescribeBuckets operation. If you're using the AWS CLI, run the describe-buckets command to retrieve these details. In your request, optionally use filter criteria to specify the name of the bucket. For more information and examples, see Filtering your S3 bucket inventory.

The following examples show how to use the AWS CLI to retrieve data sensitivity details for an S3 bucket. This first example retrieves the current sensitivity score and aggregated analysis statistics for a bucket.

```
$ aws macie2 get-resource-profile --resource-arn arn:aws:s3:::amzn-s3-demo-bucket
```

Where *arn:aws:s3:::amzn-s3-demo-bucket* is the ARN of the bucket. If the request succeeds, you receive output similar to the following:

```
{
    "profileUpdatedAt": "2024-11-21T15:44:46+00:00",
    "sensitivityScore": 83,
    "sensitivityScoreOverridden": false,
    "statistics": {
        "totalBytesClassified": 933599,
        "totalDetections": 3641,
        "totalDetectionsSuppressed": 0,
        "totalItemsClassified": 111,
        "totalItemsSensitive": 84,
        "totalItemsSkipped": 1,
```

```
        "totalItemsSkippedInvalidEncryption": 0,
        "totalItemsSkippedInvalidKms": 0,
        "totalItemsSkippedPermissionDenied": 0
    }
}
```

The next example retrieves a breakdown of the types of sensitive data that Macie has found in an S3 bucket, and the number of occurrences of each type. The breakdown also specifies which managed data identifier or custom data identifier detected the data. It also indicates whether the occurrences are currently excluded (suppressed) from the bucket's sensitivity score, if the score is calculated automatically by Macie.

```
$ aws macie2 list-resource-profile-detections --resource-arn arn:aws:s3:::amzn-s3-demo-bucket
```

Where *arn:aws:s3:::amzn-s3-demo-bucket* is the ARN of the bucket. If the request succeeds, you receive output similar to the following:

```
{
    "detections": [
        {
            "count": 8,
            "id": "AWS_CREDENTIALS",
            "name": "AWS_CREDENTIALS",
            "suppressed": false,
            "type": "MANAGED"
        },
        {
            "count": 1194,
            "id": "CREDIT_CARD_NUMBER",
            "name": "CREDIT_CARD_NUMBER",
            "suppressed": false,
            "type": "MANAGED"
        },
        {
            "count": 1194,
            "id": "CREDIT_CARD_SECURITY_CODE",
            "name": "CREDIT_CARD_SECURITY_CODE",
            "suppressed": false,
            "type": "MANAGED"
        },
        {
```

```
            "arn": "arn:aws:macie2:us-east-1:123456789012:custom-data-
identifier/3293a69d-4a1e-4a07-8715-208ddexample",
            "count": 8,
            "id": "3293a69d-4a1e-4a07-8715-208ddexample",
            "name": "Employee IDs with keyword",
            "suppressed": false,
            "type": "CUSTOM"
        },
        {
            "count": 1237,
            "id": "USA_SOCIAL_SECURITY_NUMBER",
            "name": "USA_SOCIAL_SECURITY_NUMBER",
            "suppressed": false,
            "type": "MANAGED"
        }
    ]
}
```

This example retrieves a list of objects that Macie selected from an S3 bucket for analysis. For each object, the list also indicates whether Macie completed its analysis of the object, and whether Macie found sensitive data in the object.

```
$ aws macie2 list-resource-profile-artifacts --resource-arn arn:aws:s3:::amzn-s3-
demo-bucket
```

Where *arn:aws:s3:::amzn-s3-demo-bucket* is the ARN of the bucket. If the request succeeds, you receive output similar to the following:

```
{
    "artifacts": [
        {
            "arn": "arn:aws:s3:::amzn-s3-demo-bucket/amzn-s3-demo-object1.csv",
            "classificationResultStatus": "COMPLETE",
            "sensitive": true
        },
        {
            "arn": "arn:aws:s3:::amzn-s3-demo-bucket/amzn-s3-demo-object2.xlsx",
            "classificationResultStatus": "COMPLETE",
            "sensitive": true
        },
        {
            "arn": "arn:aws:s3:::amzn-s3-demo-bucket/amzn-s3-demo-object3.json",
```

```
            "classificationResultStatus": "COMPLETE",
            "sensitive": true
        },
        {
            "arn": "arn:aws:s3:::amzn-s3-demo-bucket/amzn-s3-demo-object4.pdf",
            "classificationResultStatus": "COMPLETE",
            "sensitive": true
        },
        {
            "arn": "arn:aws:s3:::amzn-s3-demo-bucket/amzn-s3-demo-object5.zip",
            "classificationResultStatus": "PARTIAL",
            "sensitive": true
        },
        {
            "arn": "arn:aws:s3:::amzn-s3-demo-bucket/amzn-s3-demo-object6.vssx",
            "classificationResultStatus": "SKIPPED"
        }
    ]
}
```


Analyzing findings from automated sensitive data discovery

When Amazon Macie performs automated sensitive data discovery, it creates a sensitive data finding for each Amazon Simple Storage Service (Amazon S3) object that it finds sensitive data in. A *sensitive data finding* is a detailed report of sensitive data that Macie found in an S3 object. A finding doesn't include the sensitive data that Macie found. Instead, it provides information that you can use for further investigation and remediation as necessary.

Each sensitive data finding provides a severity rating and details such as:

- The date and time when Macie found the sensitive data.
- The category and types of sensitive data that Macie found.
- The number of occurrences of each type of sensitive data that Macie found.
- How Macie found the sensitive data, automated sensitive data discovery or a sensitive data discovery job.
- The name, public access settings, encryption type, and other information about the affected S3 bucket and object.

Depending on the affected S3 object's file type or storage format, the details can also include the location of as many as 15 occurrences of the sensitive data that Macie found.

Macie stores sensitive data findings for 90 days. You can access them by using the Amazon Macie console or the Amazon Macie API. You can also monitor and process findings by using other applications, services, and systems. For more information, see Reviewing and analyzing findings.

To analyze findings produced by automated sensitive data discovery

To identify and analyze findings that Macie created while performing automated sensitive data discovery, you can filter your findings. With filters, you use specific attributes of findings to build custom views and queries for findings. To filter findings, you can use the Amazon Macie console or submit queries programmatically using the Amazon Macie API. For more information, see Filtering findings.

> **Note**
>
> If your account is part of an organization that centrally manages multiple Macie accounts, only the Macie administrator for your organization has direct access to findings that automated sensitive data discovery produces for accounts in your organization. If you have a member account and want to review the findings for your account, contact your Macie administrator.

Console

Follow these steps to identify and analyze the findings by using the Amazon Macie console.

To analyze findings produced by automated discovery

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Findings**.
3. To display findings that were suppressed by a suppression rule, change the **Finding status** setting. Choose **All** to display both suppressed and unsuppressed findings, or choose **Archived** to display only suppressed findings. To then hide suppressed findings again, choose **Current**.
4. Place your cursor in the **Filter criteria** box. In the list of fields that appears, choose **Origin type**.

This field specifies how Macie found the sensitive data that produced a finding, automated sensitive data discovery or a sensitive data discovery job. To find this field in the list of filter fields, you can browse the complete list, or enter part of the field's name to narrow the list of fields.

5. Select **AUTOMATED_SENSITIVE_DATA_DISCOVERY** as the value for the field, and then choose **Apply**. Macie applies the filter criteria and adds the condition to a filter token in the **Filter criteria** box.
6. To refine the results, add filter conditions for additional fields—for example, **Created at** for the time range when a finding was created, **S3 bucket name** for the name of an affected bucket, or **Sensitive data detection type** for the type of sensitive that was detected and produced a finding.

If you want to subsequently use this set of conditions again, you can save it as a filter rule. To do this, choose **Save rule** in the **Filter criteria** box. Then enter a name and, optionally, a description for the rule. When you finish, choose **Save**.

API

To identify and analyze the findings programmatically, specify filter criteria in queries that you submit using the ListFindings or GetFindingStatistics operation of the Amazon Macie API. The **ListFindings** operation returns an array of finding IDs, one ID for each finding that matches the filter criteria. You can then use those IDs to retrieve the details of each finding. The **GetFindingStatistics** operation returns aggregated statistical data about all the findings that match the filter criteria, grouped by a field that you specify in your request. For more information about filtering findings programmatically, see Filtering findings.

In the filter criteria, include a condition for the `originType` field. This field specifies how Macie found the sensitive data that produced a finding, automated sensitive data discovery or a sensitive data discovery job. If automated sensitive data discovery produced a finding, the value for this field is `AUTOMATED_SENSITIVE_DATA_DISCOVERY`.

To identify and analyze the findings by using the AWS Command Line Interface (AWS CLI), run the list-findings or get-finding-statistics command. The following examples use the **list-findings** command to retrieve finding IDs for all high-severity findings that automated sensitive data discovery produced in the current AWS Region.

This example is formatted for Linux, macOS, or Unix, and it uses the backslash (\) line-continuation character to improve readability.

```
$ aws macie2 list-findings \
--finding-criteria '{"criterion":{"classificationDetails.originType":{"eq":
["AUTOMATED_SENSITIVE_DATA_DISCOVERY"]},"severity.description":{"eq":["High"]}}}'
```

This example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws macie2 list-findings ^
--finding-criteria={\"criterion\":{\"classificationDetails.originType\":{\"eq
\":[\"AUTOMATED_SENSITIVE_DATA_DISCOVERY\"]},\"severity.description\":{\"eq\":
[\"High\"]}}}
```

Where:

- `classificationDetails.originType` specifies the JSON name of the **Origin type** field, and:
  - `eq` specifies the *equals* operator.
  - `AUTOMATED_SENSITIVE_DATA_DISCOVERY` is an enumerated value for the field.
- *`severity.description`* specifies the JSON name of the **Severity** field, and:
  - *`eq`* specifies the *equals* operator.
  - *`High`* is an enumerated value for the field.

If the request succeeds, Macie returns a `findingIds` array. The array lists the unique identifier for each finding that matches the filter criteria, as shown in the following example.

```
{
    "findingIds": [
        "1f1c2d74db5d8caa76859ec52example",
        "6cfa9ac820dd6d55cad30d851example",
        "702a6fd8750e567d1a3a63138example",
        "826e94e2a820312f9f964cf60example",
        "274511c3fdcd87010a19a3a42example"
    ]
}
```

If no findings match the filter criteria, Macie returns an empty `findingIds` array.

```
{
```

```
    "findingIds": []
}
```

Accessing discovery results from automated sensitive data discovery

When Amazon Macie performs automated sensitive data discovery, it creates an analysis record for each Amazon Simple Storage Service (Amazon S3) object that it selects for analysis. These records, referred to as *sensitive data discovery results*, log details about the analysis that Macie performs on individual S3 objects. This includes objects that Macie doesn't find sensitive data in, and objects that Macie can't analyze due to errors or issues such as permissions settings or use of an unsupported file or storage format. Sensitive data discovery results provide you with analysis records that can be helpful for data privacy and protection audits or investigations.

If Macie finds sensitive data in an S3 object, the sensitive data discovery result provides information about the sensitive data that Macie found. The information includes the same types of details that a sensitive data finding provides. It provides additional information too, such as the location of as many as 1,000 occurrences of each type of sensitive data that Macie found. For example:

- The column and row number for a cell or field in a Microsoft Excel workbook, CSV file, or TSV file
- The path to a field or array in a JSON or JSON Lines file
- The line number for a line in a non-binary text file other than a CSV, JSON, JSON Lines, or TSV file—for example, an HTML, TXT, or XML file
- The page number for a page in an Adobe Portable Document Format (PDF) file
- The record index and the path to a field in a record in an Apache Avro object container or Apache Parquet file

If the affected S3 object is an archive file, such as a .tar or .zip file, the sensitive data discovery result also provides detailed location data for occurrences of sensitive data in individual files that Macie extracted from the archive. Macie doesn't include this information in sensitive data findings for archive files. To report location data, sensitive data discovery results use a standardized JSON schema.

> **Note**
>
> As is the case with sensitive data findings, sensitive data discovery results don't include sensitive data that Macie finds in S3 objects. Instead, they provide analysis details that can be helpful for audits or investigations.

Macie stores your sensitive data discovery results for 90 days. You can't access them directly on the Amazon Macie console or with the Amazon Macie API. Instead, you configure Macie to encrypt and store them in an S3 bucket. The bucket can serve as a definitive, long-term repository for all of your sensitive data discovery results. To determine where this repository is for your account, choose **Discovery results** in the navigation pane on the Amazon Macie console. To do this programmatically, use the GetClassificationExportConfiguration operation of the Amazon Macie API. If you haven't configured this repository for your account, see Storing and retaining sensitive data discovery results to learn how.

After you configure Macie to store your sensitive data discovery results in an S3 bucket, Macie writes the results to JSON Lines (.jsonl) files, and it encrypts and adds those files to the bucket as GNU Zip (.gz) files. For automated sensitive data discovery, Macie adds the files to a folder named `automated-sensitive-data-discovery` in the bucket. You can then optionally access and query the results in that folder. If your account is part of an organization that centrally manages multiple Macie accounts, Macie adds the files to the `automated-sensitive-data-discovery` folder in the bucket for your Macie administrator's account.

Sensitive data discovery results adhere to a standardized schema. This can help you query, monitor, and process them by using other applications, services, and systems. For a detailed, instructional example of how you might query and use these results, see the following blog post on the *AWS Security Blog*: How to query and visualize Macie sensitive data discovery results with Amazon Athena and Amazon QuickSight. For samples of Athena queries that you can use to analyze the results, visit the Amazon Macie Results Analytics repository on GitHub. This repository also provides instructions for configuring Athena to retrieve and decrypt your results, and scripts for creating tables for the results.

Assessing automated sensitive data discovery coverage

As automated sensitive data discovery progresses for your account or organization, Amazon Macie provides statistics and details to help you assess and monitor its coverage of your Amazon Simple Storage Service (Amazon S3) data estate. With this data, you can check the status of

automated sensitive data discovery for your data estate overall and individual S3 buckets within it. You can also identify issues that prevented Macie from analyzing objects in specific buckets. If you remediate the issues, you can increase coverage of your Amazon S3 data during subsequent analysis cycles.

Coverage data provides a snapshot of the current status of automated sensitive data discovery for your S3 general purpose buckets in the current AWS Region. If you're the Macie administrator for an organization, this includes buckets that your member accounts own. For each bucket, the data indicates whether issues occurred when Macie attempted to analyze objects in the bucket. If issues occurred, the data indicates the nature of each issue and, in certain cases, the number of occurrences. The data is updated as automated sensitive data discovery progresses each day. If Macie analyzes or attempts to analyze one or more objects in a bucket during a daily analysis cycle, Macie updates coverage and other data to reflect the results.

For certain types of issues, you can review the data in aggregate for all of your S3 general purpose buckets and optionally drill down for additional details about each bucket. For example, coverage data can help you quickly identify all the buckets that Macie isn't allowed to access for your account. Coverage data also reports object-level issues that occurred. These issues, referred to as *classification errors*, prevented Macie from analyzing specific objects in a bucket. For example, you can determine how many objects Macie couldn't analyze in a bucket because the objects are encrypted with an AWS Key Management Service (AWS KMS) key that's no longer available.

If you use the Amazon Macie console to review coverage data, your view of the data includes guidance for remediating each type of issue. Subsequent topics in this section also provide remediation guidance for each type.

Topics

Reviewing coverage data for automated sensitive data discovery

To review and assess coverage by automated sensitive data discovery, you can use the Amazon Macie console or the Amazon Macie API. Both the console and the API provide data that indicates the current status of the analyses for your Amazon Simple Storage Service (Amazon S3) general purpose buckets in the current AWS Region. The data includes information about issues that create gaps in the analyses:

- Buckets that Macie isn't allowed to access. Macie can't analyze any objects in these buckets. The buckets' permissions settings prevent Macie from accessing the buckets and the buckets' objects.
- Buckets that don't store any classifiable objects. Macie can't analyze any objects in these buckets. All the objects use Amazon S3 storage classes that Macie doesn't support, or they have file name extensions for file or storage formats that Macie doesn't support.
- Buckets that Macie hasn't been able to analyze yet due to object-level classification errors. Macie attempted to analyze one or more objects in these buckets. However, Macie couldn't analyze the objects due to issues with object-level permissions settings, object content, or quotas.

Coverage data is updated as automated sensitive data discovery progresses each day. If you're the Macie administrator for an organization, the data includes information for S3 buckets that your member accounts own.

> **Note**
>
> Coverage data doesn't explicitly include results for sensitive data discovery jobs that you create and run. However, remediating coverage issues that affect automated sensitive data discovery is likely to also increase coverage by jobs that you subsequently run. To assess coverage for a job, review the job's results. If a job's log events or other results indicate coverage issues, remediation guidance for automated sensitive data discovery can help you address some of the issues.

To review coverage data for automated sensitive data discovery

To review coverage data for automated sensitive data discovery, you can use the Amazon Macie console or the Amazon Macie API. On the console, a single page provides a unified view of coverage data for all of your S3 general purpose buckets in the current Region. This includes a rollup of issues that recently occurred for each bucket. The page also provides options for reviewing groups of data by issue type. To track your investigation of issues for specific buckets, you can export data from the page to a comma-separated values (CSV) file.

Console

Follow these steps to review coverage data by using the Amazon Macie console.

To review coverage data

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Resource coverage**.
3. On the **Resource coverage** page, choose the tab for the type of coverage data that you want to review:

   - **All** – Lists all the buckets for your account. For each bucket, the **Issues** field indicates whether issues prevented Macie from analyzing objects in the bucket. If the value for this field is **None**, Macie has analyzed at least one of the bucket's objects or Macie hasn't attempted to analyze any of the bucket's objects yet. If there are issues, this field indicates the nature of the issues and how to remediate them. For object-level classification errors, it might also indicate (in parentheses) the number of occurrences of the error.
   - **Access denied** – Lists buckets that Macie isn't allowed to access. The permissions settings for these buckets prevent Macie from accessing the buckets and the buckets' objects. Consequently, Macie can't analyze any objects in the buckets.
   - **Classification error** – Lists buckets that Macie hasn't analyzed yet due to object-level classification errors—issues with object-level permissions settings, object content, or quotas. For each bucket, the **Issues** field indicates the nature of each type of error that occurred and prevented Macie from analyzing an object in the bucket. It also indicates how to remediate each type of error. Depending on the error, it might also indicate (in parentheses) the number of occurrences of the error.
   - **Unclassifiable** – Lists buckets that Macie can't analyze because they don't store any classifiable objects. All the objects in these buckets use unsupported Amazon S3 storage classes or they have file name extensions for unsupported file or storage formats. Consequently, Macie can't analyze any objects in the buckets.

4. To drill down and review the supporting data for a bucket, choose the bucket's name. Then refer to the details panel for statistics and other information about the bucket.
5. To export the table to a CSV file, choose **Export to CSV** at the top of the page. The resulting CSV file contains a subset of metadata for each bucket in the table, for up to 50,000 buckets. The file includes a **Coverage issues** field. The value for this field indicates whether issues prevented Macie from analyzing objects in the bucket and, if so, the nature of the issues.

API

To review coverage data programmatically, specify filter criteria in queries that you submit using the DescribeBuckets operation of the Amazon Macie API. This operation returns an array of objects. Each object contains statistical data and other information about an S3 general purpose bucket that matches the filter criteria.

In the filter criteria, include a condition for the type of coverage data that you want to review:

- To identify buckets that Macie isn't allowed to access due to the buckets' permissions settings, include a condition where the value for the `errorCode` field equals `ACCESS_DENIED`.
- To identify buckets that Macie is allowed to access and hasn't analyzed yet, include conditions where the value for the `sensitivityScore` field equals `50` and the value for the `errorCode` field doesn't equal `ACCESS_DENIED`.
- To identify buckets that Macie can't analyze because all the buckets' objects use unsupported storage classes or formats, include conditions where the value for the `classifiableSizeInBytes` field equals `0` and the value for the `sizeInBytes` field is greater than `0`.
- To identify buckets for which Macie has analyzed at least one object, include conditions where the value for the `sensitivityScore` field falls within the range of 1–99 but is not equal to `50`. To also include buckets where you manually assigned the maximum score, the range should be 1–100.
- To identify buckets that Macie hasn't analyzed yet due to object-level classification errors, include a condition where the value for the `sensitivityScore` field equals `-1`. To then review a breakdown of the types and number of errors that occurred for a particular bucket, use the GetResourceProfile operation.

If you're using the AWS Command Line Interface (AWS CLI), specify filter criteria in queries that you submit by running the describe-buckets command. To review a breakdown of the types and number of errors that occurred for a particular S3 bucket, if any, run the get-resource-profile command.

For example, the following AWS CLI commands use filter criteria to retrieve the details of all the S3 buckets that Macie isn't allowed to access due to the buckets' permissions settings.

This example is formatted for Linux, macOS, or Unix:

```
$ aws macie2 describe-buckets --criteria '{"errorCode":{"eq":["ACCESS_DENIED"]}}'
```

This example is formatted for Microsoft Windows:

```
C:\> aws macie2 describe-buckets --criteria={\"errorCode\":{\"eq\":[\"ACCESS_DENIED
\"]}}
```

If your request succeeds, Macie returns a `buckets` array. The array contains an object for each S3 bucket that's in the current AWS Region and matches the filter criteria.

If no S3 buckets match the filter criteria, Macie returns an empty `buckets` array.

```
{
    "buckets": []
}
```

For more information about specifying filter criteria in queries, including examples of common criteria, see Filtering your S3 bucket inventory.

For detailed information that can help you address coverage issues, see Remediating coverage issues for automated sensitive data discovery.

Remediating coverage issues for automated sensitive data discovery

As automated sensitive data discovery progresses each day, Amazon Macie provides statistics and details to help you assess and monitor its coverage of your Amazon Simple Storage Service (Amazon S3) data estate. By reviewing coverage data, you can check the status of automated sensitive data discovery for your data estate overall and individual S3 buckets within it. You can also identify issues that prevented Macie from analyzing objects in specific buckets. If you remediate the issues, you can increase coverage of your Amazon S3 data during subsequent analysis cycles.

Macie reports several types of issues that reduce coverage of your Amazon S3 data by automated sensitive data discovery. This includes bucket-level issues that prevent Macie from analyzing any objects in an S3 bucket. It also includes object-level issues. These issues, referred to as *classification errors*, prevented Macie from analyzing specific objects in a bucket. The following information can help you investigate and remediate the issues.

Issue types and details

- Access denied
- Classification error: Invalid content
- Classification error: Invalid encryption
- Classification error: Invalid KMS key
- Classification error: Permission denied
- Unclassifiable

> **Tip**
>
> To investigate object-level classification errors for an S3 bucket, start by reviewing the list of object samples for the bucket. This list indicates which objects Macie analyzed or attempted to analyze in the bucket, for up to 100 objects.
> To review the list on the Amazon Macie console, choose the bucket on the **S3 buckets** page, and then choose the **Object samples** tab in the details panel. To review the list programmatically, use the ListResourceProfileArtifacts operation of the Amazon Macie API. If the status of the analysis for an object is **Skipped** (`SKIPPED`), the object might have caused the error.

Access denied

This issue indicates that an S3 bucket's permissions settings prevent Macie from accessing the bucket and the bucket's objects. Macie can't retrieve and analyze any objects in the bucket.

Details

The most common cause for this type of issue is a restrictive bucket policy. A *bucket policy* is a resource-based AWS Identity and Access Management (IAM) policy that specifies which actions a principal (user, account, service, or other entity) can perform on an S3 bucket, and the conditions under which a principal can perform those actions. A *restrictive bucket policy* uses explicit `Allow` or `Deny` statements that grant or restrict access to a bucket's data based on specific conditions. For example, a bucket policy might contain an `Allow` or `Deny` statement that denies access to a bucket unless specific source IP addresses are used to access the bucket.

If the bucket policy for an S3 bucket contains an explicit `Deny` statement with one or more conditions, Macie might not be allowed to retrieve and analyze the bucket's objects to detect

sensitive data. Macie can only provide a subset of information about the bucket, such as the bucket's name and creation date.

Remediation guidance

To remediate this issue, update the bucket policy for the S3 bucket. Ensure that the policy allows Macie to access the bucket and the bucket's objects. To allow this access, add a condition for the Macie service-linked role (`AWSServiceRoleForAmazonMacie`) to the policy. The condition should exclude the Macie service-linked role from matching the `Deny` restriction in the policy. It can do this by using the `aws:PrincipalArn` global condition context key and the Amazon Resource Name (ARN) of the Macie service-linked role for your account.

If you update the bucket policy and Macie gains access to the S3 bucket, Macie will detect the change. When this happens, Macie will update statistics, inventory data, and other information that it provides about your Amazon S3 data. In addition, the bucket's objects will be a higher priority for analysis during a subsequent analysis cycle.

Additional reference

For more information about updating an S3 bucket policy to allow Macie to access a bucket, see Allowing Macie to access S3 buckets and objects. For information about using bucket policies to control access to buckets, see Bucket policies and How Amazon S3 authorizes a request in the *Amazon Simple Storage Service User Guide*.

Classification error: Invalid content

This type of classification error occurs if Macie attempts to analyze an object in an S3 bucket and the object is malformed or the object contains content that exceeds a sensitive data discovery quota. Macie can't analyze the object.

Details

This error typically occurs because an S3 object is a malformed or corrupted file. Consequently, Macie can't parse and analyze all the data in the file.

This error can also occur if analysis of an S3 object would exceed a sensitive data discovery quota for an individual file. For example, the storage size of the object exceeds the size quota for that type of file.

For either case, Macie can't complete its analysis of the S3 object and the status of the analysis for the object is **Skipped** (`SKIPPED`).

Remediation guidance

To investigate this error, download the S3 object and check the formatting and contents of the file. Also assess the contents of the file against Macie quotas for sensitive data discovery.

If you don't remediate this error, Macie will try to analyze other objects in the S3 bucket. If Macie analyzes another object successfully, Macie will update coverage data and other information that it provides about the bucket.

Additional reference

For a list of sensitive data discovery quotas, including the quotas for certain types of files, see Quotas for Macie. For information about how Macie updates sensitivity scores and other information that it provides about S3 buckets, see How automated sensitive data discovery works.

Classification error: Invalid encryption

This type of classification error occurs if Macie attempts to analyze an object in an S3 bucket and the object is encrypted with a customer-provided key. The object uses SSE-C encryption, which means that Macie can't retrieve and analyze the object.

Details

Amazon S3 supports multiple encryption options for S3 objects. For most of these options, Macie can decrypt an object by using the Macie service-linked role for your account. However, this depends on the type of encryption that was used.

For Macie to decrypt an S3 object, the object must be encrypted with a key that Macie can access and is allowed to use. If an object is encrypted with a customer-provided key, Macie can't provide the requisite key material to retrieve the object from Amazon S3. Consequently, Macie can't analyze the object and the status of the analysis for the object is **Skipped** (`SKIPPED`).

Remediation guidance

To remediate this error, encrypt S3 objects with Amazon S3 managed keys or AWS Key Management Service (AWS KMS) keys. If you prefer to use AWS KMS keys, the keys can be AWS managed KMS keys, or customer managed KMS keys that Macie is allowed to use.

To encrypt existing S3 objects with keys that Macie can access and use, you can change the encryption settings for the objects. To encrypt new objects with keys that Macie can access and

use, change the default encryption settings for the S3 bucket. Also ensure that the bucket's policy doesn't require new objects to be encrypted with a customer-provided key.

If you don't remediate this error, Macie will try to analyze other objects in the S3 bucket. If Macie analyzes another object successfully, Macie will update coverage data and other information that it provides about the bucket.

Additional reference

For information about requirements and options for using Macie to analyze encrypted S3 objects, see Analyzing encrypted Amazon S3 objects. For information about encryption options and settings for S3 buckets, see Protecting data with encryption and Setting default server-side encryption behavior for S3 buckets in the *Amazon Simple Storage Service User Guide*.

Classification error: Invalid KMS key

This type of classification error occurs if Macie attempts to analyze an object in an S3 bucket and the object is encrypted with an AWS Key Management Service (AWS KMS) key that's no longer available. Macie can't retrieve and analyze the object.

Details

AWS KMS provides options for disabling and deleting customer managed AWS KMS keys. If an S3 object is encrypted with a KMS key that is disabled, is scheduled for deletion, or was deleted, Macie can't retrieve and decrypt the object. Consequently, Macie can't analyze the object and the status of the analysis for the object is **Skipped** (`SKIPPED`). For Macie to analyze an encrypted object, the object must be encrypted with a key that Macie can access and is allowed to use.

Remediation guidance

To remediate this error, re-enable the applicable AWS KMS key or cancel the scheduled deletion of the key, depending on the current status of the key. If the applicable key was already deleted, this error cannot be remediated.

To determine which AWS KMS key was used to encrypt an S3 object, you can start by using Macie to review the server-side encryption settings for the S3 bucket. If the default encryption settings for the bucket are configured to use a KMS key, the bucket's details indicate which key is used. You can then check the status of that key. Alternatively, you can use Amazon S3 to review the encryption settings for the bucket and individual objects in the bucket.

If you don't remediate this error, Macie will try to analyze other objects in the S3 bucket. If Macie analyzes another object successfully, Macie will update coverage data and other information that it provides about the bucket.

Additional reference

For information about using Macie to review the server-side encryption settings for an S3 bucket, see Reviewing the details of S3 buckets. For information about re-enabling an AWS KMS key or canceling the scheduled deletion of a key, see Enabling and disabling keys and Deleting keys in the *AWS Key Management Service Developer Guide*.

Classification error: Permission denied

This type of classification error occurs if Macie attempts to analyze an object in an S3 bucket and Macie can't retrieve or decrypt the object due to the permissions settings for the object or the permissions settings for the key that was used to encrypt the object. Macie can't retrieve and analyze the object.

Details

This error typically occurs because an S3 object is encrypted with a customer managed AWS Key Management Service (AWS KMS) key that Macie isn't allowed to use. If an object is encrypted with a customer managed AWS KMS key, the key's policy must allow Macie to decrypt data by using the key.

This error can also occur if Amazon S3 permissions settings prevent Macie from retrieving an S3 object. The bucket policy for the S3 bucket might restrict access to specific bucket objects or allow only certain principals (users, accounts, services, or other entities) to access the objects. Or the access control list (ACL) for an object might restrict access to the object. Consequently, Macie might not be allowed to access the object.

For any of the preceding cases, Macie can't retrieve and analyze the object, and the status of the analysis for the object is **Skipped** (`SKIPPED`).

Remediation guidance

To remediate this error, determine whether the S3 object is encrypted with a customer managed AWS KMS key. If it is, ensure that the key's policy allows the Macie service-linked role (`AWSServiceRoleForAmazonMacie`) to decrypt data with the key. How you allow this access depends on whether the account that owns the AWS KMS key also owns the S3 bucket that stores the object. If the same account owns the KMS key and the bucket, a user of the account

has to update the key's policy. If one account owns the KMS key and a different account owns the bucket, a user of the account that owns the key has to allow cross-account access to the key.

> **Tip**
>
> You can automatically generate a list of all the customer managed AWS KMS keys that Macie needs to access to analyze objects in the S3 buckets for your account. To do this, run the AWS KMS Permission Analyzer script, which is available from the Amazon Macie Scripts repository on GitHub. The script can also generate an additional script of AWS Command Line Interface (AWS CLI) commands. You can optionally run those commands to update the requisite configuration settings and policies for KMS keys that you specify.

If Macie is already allowed to use the applicable AWS KMS key or the S3 object isn't encrypted with a customer managed KMS key, ensure that the bucket's policy allows Macie to access the object. Also verify that the object's ACL allows Macie to read the object's data and metadata.

For the bucket policy, you can allow this access by adding a condition for the Macie service-linked role to the policy. The condition should exclude the Macie service-linked role from matching the `Deny` restriction in the policy. It can do this by using the `aws:PrincipalArn` global condition context key and the Amazon Resource Name (ARN) of the Macie service-linked role for your account.

For the object ACL, you can allow this access by working with the object owner to add your AWS account as a grantee with `READ` permissions for the object. Macie can then use the service-linked role for your account to retrieve and analyze the object. Also consider changing the Object Ownership settings for the bucket. You can use these settings to disable ACLs for all the objects in the bucket and grant ownership permissions to the account that owns the bucket.

If you don't remediate this error, Macie will try to analyze other objects in the S3 bucket. If Macie analyzes another object successfully, Macie will update coverage data and other information that it provides about the bucket.

Additional reference

For more information about allowing Macie to decrypt data with a customer managed AWS KMS key, see Allowing Macie to use a customer managed AWS KMS key. For information about updating an S3 bucket policy to allow Macie to access a bucket, see Allowing Macie to access S3 buckets and objects.

For information about updating a key policy, see Changing a key policy in the *AWS Key Management Service Developer Guide*. For information about using customer managed AWS KMS keys to encrypt S3 objects, see Using server-side encryption with AWS KMS keys in the *Amazon Simple Storage Service User Guide*.

For information about using bucket policies to control access to S3 buckets, see Access control and How Amazon S3 authorizes a request in the *Amazon Simple Storage Service User Guide*. For information about using ACLs or Object Ownership settings to control access to S3 objects, see Managing access with ACLs and Controlling ownership of objects and disabling ACLs for your bucket in the *Amazon Simple Storage Service User Guide*.

Unclassifiable

This issue indicates that all the objects in an S3 bucket are stored using unsupported Amazon S3 storage classes or unsupported file or storage formats. Macie can't analyze any objects in the bucket.

Details

To be eligible for selection and analysis, an S3 object must use an Amazon S3 storage class that Macie supports. The object must also have a file name extension for a file or storage format that Macie supports. If an object doesn't meet these criteria, the object is treated as an *unclassifiable object*. Macie doesn't attempt to retrieve or analyze data in unclassifiable objects.

If all the objects in an S3 bucket are unclassifiable objects, the overall bucket is an *unclassifiable bucket*. Macie can't perform automated sensitive data discovery for the bucket.

Remediation guidance

To address this issue, review lifecycle configuration rules and other settings that determine which storage classes are used to store objects in the S3 bucket. Consider adjusting those settings to use storage classes that Macie supports. You can also change the storage class of existing objects in the bucket.

Also assess the file and storage formats of existing objects in the S3 bucket. To analyze the objects, consider porting the data, either temporarily or permanently, to new objects that use a supported format.

If objects are added to the S3 bucket and they use a supported storage class and format, Macie will detect the objects the next time it evaluates your bucket inventory. When this happens, Macie will stop reporting that the bucket is *unclassifiable* in statistics, coverage data, and other

information that it provides about your Amazon S3 data. In addition, the new objects will be a higher priority for analysis during a subsequent analysis cycle.

Additional reference

For information about the Amazon S3 storage classes and the file and storage formats that Macie supports, see Supported storage classes and formats. For information about lifecycle configuration rules and the storage class options that Amazon S3 provides, see Managing your storage lifecycle and Using Amazon S3 storage classes in the *Amazon Simple Storage Service User Guide*.

Adjusting sensitivity scores for S3 buckets

As you review and evaluate statistics, data, and other results of automated sensitive data discovery, there might be cases where you want to fine tune sensitivity assessments of your Amazon Simple Storage Service (Amazon S3) buckets. You might also want to capture the results of investigations that you or your organization performs for specific buckets. If you're the Amazon Macie administrator for an organization or you have a standalone Macie account, you can make these changes by adjusting the sensitivity score and other settings for individual buckets. If you have a member account in an organization, work with your Macie administrator to adjust the settings for buckets that you own. Only the Macie administrator for your organization can adjust these settings for your buckets.

If you're a Macie administrator or you have a standalone Macie account, you can adjust the sensitivity score for an S3 bucket in the following ways:

- **Assign a sensitivity score** – By default, Macie automatically calculates a bucket's sensitivity score. The score is based primarily on the amount of sensitive data that Macie has found in a bucket, and the amount of data that Macie has analyzed in a bucket. For more information, see Sensitivity scoring for S3 buckets.

  You can override a bucket's calculated score and manually assign the maximum score (*100*), which also applies the *Sensitive* label to the bucket. If you do this, Macie continues to perform automated sensitive data discovery for the bucket. However, subsequent analyses don't affect the bucket's score. To calculate the score automatically again, change the setting again.
- **Exclude or include sensitive data types in the sensitivity score** – If it's calculated automatically, a bucket's sensitivity score is based partly on the amount of sensitive data that Macie has found in the bucket. This derives primarily from the nature and number of sensitive data types

that Macie has found, and the number of occurrences of each type. By default, Macie includes occurrences of all types of sensitive data when it calculates a bucket's score.

You can adjust the calculation by excluding or including specific types of sensitive data in a bucket's score. For example, if Macie detected mailing addresses in a bucket and you determine that this is acceptable, you can exclude all occurrences of mailing addresses from the bucket's score. If you exclude a sensitive data type, Macie continues to inspect the bucket for that type of data, and report occurrences that it finds. However, those occurrences don't affect the bucket's score. To include a sensitive data type in the score again, change the setting again.

You can also exclude an S3 bucket from subsequent analyses. If you exclude a bucket, existing sensitive data discovery statistics and details for the bucket persist. For example, the bucket's current sensitivity score remains unchanged. However, Macie stops analyzing objects in the bucket when it performs automated sensitive data discovery. After you exclude a bucket, you can include it again later.

If you change a setting that affects the sensitivity score for an S3 bucket, Macie immediately begins to recalculate the score. Macie also updates relevant statistics and other information that it provides about the bucket and your Amazon S3 data overall. For example, if you assign the maximum score to a bucket, Macie increments the count of *Sensitive* buckets in aggregated statistics.

To adjust the sensitivity score or other settings for an S3 bucket

To adjust the sensitivity score or other settings for an S3 bucket, you can use the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to adjust the sensitivity score or a setting for an S3 bucket by using the Amazon Macie console.

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **S3 buckets**. The **S3 buckets** page displays your bucket inventory.

   By default, the page doesn't display data for buckets that are currently excluded from analyses. If you're the Macie administrator for an organization, it also doesn't display data for accounts that automated sensitive data discovery is currently disabled for. To display

this data, choose **X** in the **Is monitored by automated discovery** filter token below the filter box.

3. Choose the S3 bucket that has a setting to adjust. You can choose the bucket by using the table view () or the interactive map ().
4. In the details panel, do any of the following:

   - To override the calculated sensitivity score and manually assign a score, turn on **Assign maximum score** (). This changes the bucket's score to *100* and applies the *Sensitive* label to the bucket.
   - To assign a sensitivity score that Macie calculates automatically, turn off **Assign maximum score** ().
   - To exclude or include specific types of sensitive data in the sensitivity score, choose the **Sensitivity** tab. In the **Detections** table, select the checkbox for the sensitive data type to exclude or include. Then, on the **Actions** menu, choose **Exclude from score** to exclude the type or choose **Include in score** to include the type.

     In the table, the **Sensitive data type** field specifies the managed data identifier or custom data identifier that detected the data. For a managed data identifier, this is a unique identifier (ID) that describes the type of sensitive data that the identifier is designed to detect—for example, **USA_PASSPORT_NUMBER** for US passport numbers. For details about each managed data identifier, see Using managed data identifiers.
   - To exclude the bucket from subsequent analyses, turn on **Exclude from automated discovery** ().
   - To include the bucket in subsequent analyses, if you previously excluded it, turn off **Exclude from automated discovery** ().

API

To adjust the sensitivity score or a setting for an S3 bucket programmatically, you have several options. The appropriate option depends on what you want to adjust.

Assign a sensitivity score

To assign a sensitivity score to an S3 bucket, use the UpdateResourceProfile operation. In your request, use the `resourceArn` parameter to specify the Amazon Resource Name (ARN) of the bucket. For the `sensitivityScoreOverride` parameter, do one of the following:

- To override the calculated score and manually assign the maximum score, specify `100`.
- To assign a score that Macie calculates automatically, omit the parameter. If this parameter is null, Macie calculates and assigns the score.

If you're using the AWS Command Line Interface (AWS CLI), run the update-resource-profile command to assign a sensitivity score to an S3 bucket. In your request, use the `resource-arn` parameter to specify the ARN of the bucket. Omit or use the `sensitivity-score-override` parameter to specify which score to assign.

If your request succeeds, Macie assigns the specified score and returns an empty response.

Exclude or include sensitive data types in the sensitivity score

To exclude or include sensitive data types in the sensitivity score for an S3 bucket, use the UpdateResourceProfileDetections operation. When you use this operation, you overwrite the current inclusion and exclusion settings for a bucket's score. Therefore, it's a good idea to first retrieve the current settings and determine which ones you want to keep. To retrieve the current settings, use the ListResourceProfileDetections operation.

When you're ready to update the settings, use the `resourceArn` parameter to specify the ARN of the S3 bucket. For the `suppressDataIdentifiers` parameter, do one of the following:

- To exclude a sensitive data type from the bucket's score, use the `type` parameter to specify the type of data identifier that detected the data, a managed data identifier (`MANAGED`) or a custom data identifier (`CUSTOM`). Use the `id` parameter to specify the unique identifier for the managed or custom data identifier that detected the data.
- To include a sensitive data type in the bucket's score, don't specify any details for the managed or custom data identifier that detected the data.

- To include all sensitive data types in the bucket's score, don't specify any values. If the value for the `suppressDataIdentifiers` parameter is null (empty), Macie includes all types of detections when it calculates the score.

If you're using the AWS CLI, run the update-resource-profile-detections command to exclude or include sensitive data types in the sensitivity score for an S3 bucket. Use the `resource-arn` parameter to specify the ARN of the bucket. Use the `suppress-data-identifiers` parameter to specify which sensitive data types to exclude or include in the bucket's score. To first retrieve and review the current settings for the bucket, run the list-resource-profile-detections command.

If your request succeeds, Macie updates the settings and returns an empty response.

Exclude or include an S3 bucket in analyses

To exclude or subsequently include an S3 bucket in analyses, use the UpdateClassificationScope operation. Or, if you're using the AWS CLI, run the update-classification-scope command. For additional details and examples, see Excluding or including S3 buckets in automated sensitive data discovery.

The following examples show how to use the AWS CLI to adjust individual settings for an S3 bucket. This first example manually assigns the maximum sensitivity score (`100`) to a bucket. It overrides the bucket's calculated score.

```
$ aws macie2 update-resource-profile --resource-arn arn:aws:s3:::amzn-s3-demo-bucket
 --sensitivity-score-override 100
```

Where *`arn:aws:s3:::amzn-s3-demo-bucket`* is the ARN of the S3 bucket.

The next example changes the sensitivity score for an S3 bucket to a score that Macie calculates automatically. The bucket currently has a manually assigned score that overrides the calculated score. This example removes that override by omitting the `sensitivity-score-override` parameter from the request.

```
$ aws macie2 update-resource-profile --resource-arn arn:aws:s3:::amzn-s3-demo-
bucket2
```

Where *`arn:aws:s3:::amzn-s3-demo-bucket2`* is the ARN of the S3 bucket.

The following examples exclude particular types of sensitive data from the sensitivity score for an S3 bucket. This example is formatted for Linux, macOS, or Unix, and it uses the backslash (\) line-continuation character to improve readability.

```
$ aws macie2 update-resource-profile-detections \
--resource-arn arn:aws:s3:::amzn-s3-demo-bucket3 \
--suppress-data-identifiers '[{"type":"MANAGED","id":"ADDRESS"},
{"type":"CUSTOM","id":"3293a69d-4a1e-4a07-8715-208ddexample"}]'
```

This example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws macie2 update-resource-profile-detections ^
--resource-arn arn:aws:s3:::amzn-s3-demo-bucket3 ^
--suppress-data-identifiers=[{\"type\":\"MANAGED\",\"id\":\"ADDRESS\"},{\"type\":
\"CUSTOM\",\"id\":\"3293a69d-4a1e-4a07-8715-208ddexample\"}]
```

Where:

- *arn:aws:s3:::amzn-s3-demo-bucket3* is the ARN of the S3 bucket.
- *ADDRESS* is the unique identifier for the managed data identifier that detected a type of sensitive data to exclude (mailing addresses).
- *3293a69d-4a1e-4a07-8715-208ddexample* is the unique identifier for the custom data identifier that detected a type of sensitive data to exclude.

This next set of examples later includes all types of sensitive data in the sensitivity score for the S3 bucket. It overwrites the current exclusion settings for the bucket by specifying an empty (null) value for the `suppress-data-identifiers` parameter. For Linux, macOS, or Unix:

```
$ aws macie2 update-resource-profile-detections --resource-arn arn:aws:s3:::amzn-s3-
demo-bucket3 --suppress-data-identifiers '[]'
```

For Microsoft Windows:

```
C:\> aws macie2 update-resource-profile-detections --resource-arn arn:aws:s3:::amzn-
s3-demo-bucket3 --suppress-data-identifiers=[]
```

Where *arn:aws:s3:::amzn-s3-demo-bucket3* is the ARN of the S3 bucket.

Sensitivity scoring for S3 buckets

If automated sensitive data discovery is enabled, Amazon Macie automatically calculates and assigns a sensitivity score to each Amazon Simple Storage Service (Amazon S3) general purpose bucket that it monitors and analyzes for an account or organization. A *sensitivity score* is a quantitative representation of the amount of sensitive data that an S3 bucket might contain. Based on that score, Macie also assigns a sensitivity label to each bucket. A *sensitivity label* is a qualitative representation of a bucket's sensitivity score. These values can serve as reference points for determining where sensitive data might reside in your Amazon S3 data estate, and identifying and monitoring potential security risks for that data.

By default, an S3 bucket's sensitivity score and label reflect the results of automated sensitive data discovery activities that Macie has performed thus far for the bucket. They don't reflect the results of sensitive data discovery jobs that you create and run. In addition, neither the score nor the label implies or otherwise indicates the criticality or importance that a bucket or a bucket's objects might have for you or your organization. However, you can override a bucket's calculated score by manually assigning the maximum score (*100*) to the bucket. This also assigns the *Sensitive* label to the bucket. To override a calculated score, you must be the Macie administrator for the account that owns the bucket, or have a standalone Macie account.

Topics

Sensitivity scoring dimensions and ranges

If it's calculated by Amazon Macie, an S3 bucket's sensitivity score is a quantitative measure of the intersection of two primary dimensions:

- The amount of sensitive data that Macie has found in the bucket. This derives primarily from the nature and number of sensitive data types that Macie has found in the bucket, and the number of occurrences of each type.
- The amount of data that Macie has analyzed in the bucket. This derives primarily from the number of unique objects that Macie has analyzed in the bucket relative to the total number of unique objects in the bucket.

An S3 bucket's sensitivity score also determines which sensitivity label Macie assigns to the bucket. The sensitivity label is a qualitative representation of the score—for example, *Sensitive* or *Not sensitive*. On the Amazon Macie console, a bucket's sensitivity score also determines which color Macie uses to represent the bucket in data visualizations, as shown in the following image.

Sensitivity scores range from *-1* through *100*, as described in the following table. To assess inputs to an S3 bucket's score, you can refer to sensitive data discovery statistics and other details that Macie provides about the bucket.

Sensitivity score	Sensitivity label	Additional information
-1	Classification error	Macie hasn't successfully analyzed any of the bucket's objects yet due to object-level classification errors—issues with object-level permissio ns settings, object content, or quotas. When Macie tried to analyze one or more objects in the bucket, errors occurred. For example, an object is a malformed file, or an object is encrypted with a key that Macie can't access or isn't allowed to use. Coverage data for the bucket can help you investigate and remediate the errors. For more informati on, see Assessing automated

Sensitivity score	Sensitivity label	Additional information
		sensitive data discovery coverage. Macie will continue to try to analyze objects in the bucket. If Macie analyzes an object successfully, Macie will update the bucket's sensitivity score and label to reflect the results of the analysis.
1-49	Not sensitive	In this range, a higher score, such as *49*, indicates that Macie has analyzed relatively few objects in the bucket. A lower score, such as *1*, indicates that Macie has analyzed many objects in the bucket (relative to the total number of objects in the bucket) and detected relatively few types and occurrences of sensitive data in those objects. A score of *1* can also indicate that the bucket doesn't store any objects or all the objects in the bucket contain zero (0) bytes of data. Object statistics in the bucket's details can help you determine if this is the case. For more information, see Reviewing S3 bucket details.

Sensitivity score	Sensitivity label	Additional information
50	Not yet analyzed	Macie hasn't tried to analyze or analyzed any of the bucket's objects yet. Macie automatically assigns this score when automated discovery is initially enabled or a bucket is added to the bucket inventory for an account. In an organization, a bucket can also have this score if automated discovery has never been enabled for the account that owns the bucket. A score of *50* can also indicate that the bucket's permissio ns settings prevent Macie from accessing the bucket or the bucket's objects. This is typically due to a restrictive bucket policy. The bucket's details can help you determine if this is the case because Macie can provide only a subset of information about the bucket. For information about how to address this issue, see Allowing Macie to access S3 buckets and objects.

Sensitivity score	Sensitivity label	Additional information
51-99	Sensitive	In this range, a higher score, such as *99*, indicates that Macie has analyzed many objects in the bucket (relative to the total number of objects in the bucket) and detected many types and occurrenc es of sensitive data in those objects. A lower score, such as *51*, indicates that Macie has analyzed a moderate number of objects in the bucket (relative to the total number of objects in the bucket) and detected at least a few types and occurrences of sensitive data in those objects.
100	Sensitive	The score was manually assigned to the bucket, overriding the calculated score. Macie doesn't assign this score to buckets.

Monitoring sensitivity scores

When automated sensitive data discovery is initially enabled for an account, Amazon Macie automatically assigns a sensitivity score of *50* to each S3 bucket that the account owns. Macie also assigns this score to a bucket when the bucket is added to the bucket inventory for an account. Based on that score, each bucket's sensitivity label is *Not yet analyzed*. The exception is an empty bucket, which is a bucket that doesn't store any objects or all the objects in the bucket contain zero (0) bytes of data. If this is the case for a bucket, Macie assigns a score of *1* to the bucket and the bucket's sensitivity label is *Not sensitive*.

As automated sensitive data discovery progresses each day, Macie updates sensitivity scores and labels for S3 buckets to reflect the results of its analysis. For example:

- If Macie doesn't find sensitive data in an object, Macie decreases the bucket's sensitivity score and updates the sensitivity label as necessary.
- If Macie finds sensitive data in an object, Macie increases the bucket's sensitivity score and updates the sensitivity label as necessary.
- If Macie finds sensitive data in an object that's subsequently changed, Macie removes sensitive data detections for the object from the bucket's sensitivity score and updates the sensitivity label as necessary.
- If Macie finds sensitive data in an object that's subsequently deleted, Macie removes sensitive data detections for the object from the bucket's sensitivity score and updates the sensitivity label as necessary.
- If an object is added to a bucket that was previously empty and Macie finds sensitive data in the object, Macie increases the bucket's sensitivity score and updates the sensitivity label as necessary.
- If a bucket's permissions settings prevent Macie from accessing or retrieving information about the bucket or the bucket's objects, Macie changes the bucket's sensitivity score to *50* and changes the bucket's sensitivity label to *Not yet analyzed*.

Analysis results can begin to appear within 48 hours of enabling automated sensitive data discovery for an account.

If you're the Macie administrator for an organization or you have a standalone Macie account, you can adjust sensitivity scoring settings for your organization or account:

- To adjust the settings for subsequent analyses of all S3 buckets, change the settings for your account. You can start including or excluding specific managed data identifiers, custom data identifiers, or allow lists. You can also exclude specific buckets. For more information, see Configuring automated discovery settings.
- To adjust the settings for individual S3 buckets, change the settings for each bucket. You can include or exclude specific types of sensitive data from a bucket's score. You can also specify whether to assign an automatically calculated score to a bucket. For more information, see Adjusting sensitivity scores for S3 buckets.

If you disable automated sensitive data discovery, the effect varies for existing sensitivity scores and labels. If you disable it for a member account in an organization, existing scores and labels persist for S3 buckets that the account owns. If you disable it for an organization overall or a standalone Macie account, existing scores and labels persist for only 30 days. After 30 days, Macie resets scores and labels for all the buckets that the organization or account owns. If a bucket stores objects, Macie changes the score to *50* and assigns the *Not yet analyzed* label to the bucket. If a bucket is empty, Macie changes the score to *1* and assigns the *Not sensitive* label to the bucket. After this reset, Macie stops updating sensitivity scores and labels for the buckets, unless you enable automated sensitive data discovery for the organization or account again.

Default settings for automated sensitive data discovery

If automated sensitive data discovery is enabled, Amazon Macie automatically selects and analyzes sample objects from all the Amazon Simple Storage Service (Amazon S3) general purpose buckets for your account. If you're the Macie administrator for an organization, by default this includes S3 buckets that your member accounts own.

If you're a Macie administrator or you have a standalone Macie account, you can refine the scope of the analyses by excluding specific S3 buckets from automated sensitive data discovery. You can do this in two ways: by changing the settings for your account, and by changing the settings for individual buckets. As a Macie administrator, you can also enable or disable automated sensitive data discovery for individual accounts in your organization.

By default, Macie analyzes S3 objects by using only the set of managed data identifiers that we recommend for automated sensitive data discovery. Macie doesn't use any custom data identifiers or allow lists that you defined. If you're a Macie administrator or you have a standalone Macie account, you can customize the analyses by configuring Macie to use specific managed data identifiers, custom data identifiers, and allow lists. You can do this by changing the settings for your account.

For information about changing your settings, see Configuring settings for automated sensitive data discovery.

Topics

- Default managed data identifiers for automated sensitive data discovery
- Updates to the default settings for automated sensitive data discovery

Default managed data identifiers for automated sensitive data discovery

By default, Amazon Macie analyzes S3 objects by using only the set of managed data identifiers that we recommend for automated sensitive data discovery. This default set of managed data identifiers is designed to detect common categories and types of sensitive data. Based on our research, it can detect general categories and types of sensitive data while also optimizing your results by reducing noise.

The default set is dynamic. As we release new managed data identifiers, we add them to the default set if they're likely to further optimize your automated sensitive data discovery results. Over time, we might also add or remove existing managed data identifiers from the set. Removal of a managed data identifier doesn't affect existing sensitive data discovery statistics and details for your S3 buckets. For example, if we remove the managed data identifier for a type of sensitive data that Macie previously detected in a bucket, Macie continues to report those detections. If we add or remove a managed data identifier from the default set, we update this page to indicate the nature and timing of the change. For automatic alerts about these changes, you can subscribe to the RSS feed on the Macie document history page.

The following topics list the managed data identifiers that are currently in the default set, organized by sensitive data category and type. They specify the unique identifier (ID) for each managed data identifier in the set. This ID describes the type of sensitive data that a managed data identifier is designed to detect, for example: `PGP_PRIVATE_KEY` for PGP private keys and `USA_PASSPORT_NUMBER` for US passport numbers. If you change your settings for automated sensitive data discovery, you can use this ID to explicitly exclude a managed data identifier from subsequent analyses.

Topics

For details about specific managed data identifiers or a complete list of all the managed data identifiers that Macie currently provides, see Using managed data identifiers.

Credentials

To detect occurrences of credentials data in S3 objects, Macie uses the following managed data identifiers by default.

Sensitive data type	Managed data identifier ID
AWS secret access key	`AWS_CREDENTIALS`
HTTP Basic Authorization header	`HTTP_BASIC_AUTH_HEADER`
OpenSSH private key	`OPENSSH_PRIVATE_KEY`
PGP private key	`PGP_PRIVATE_KEY`
Public Key Cryptography Standard (PKCS) private key	`PKCS`
PuTTY private key	`PUTTY_PRIVATE_KEY`

Financial information

To detect occurrences of financial information in S3 objects, Macie uses the following managed data identifiers by default.

Sensitive data type	Managed data identifier ID
Credit card magnetic stripe data	`CREDIT_CARD_MAGNETIC_STRIPE`
Credit card number	`CREDIT_CARD_NUMBER` (for credit card numbers in proximity of a keyword)

Personally identifiable information (PII)

To detect occurrences of personally identifiable information (PII) in S3 objects, Macie uses the following managed data identifiers by default.

Sensitive data type	Managed data identifier ID
Driver's license identification number	`CANADA_DRIVERS_LICENSE`, `DRIVERS_LICENSE` (for the US), `UK_DRIVERS_LICENSE`

Sensitive data type	Managed data identifier ID
Electoral roll number	UK_ELECTORAL_ROLL_NUMBER
National identification number	FRANCE_NATIONAL_IDENTIFICATION_NUMBER, GERMANY_NATIONAL_IDENTIFICATION_NUMBER, ITALY_NATIONAL_IDENTIFICATION_NUMBER, SPAIN_DNI_NUMBER
National Insurance Number (NINO)	UK_NATIONAL_INSURANCE_NUMBER
Passport number	CANADA_PASSPORT_NUMBER, FRANCE_PASSPORT_NUMBER, GERMANY_PASSPORT_NUMBER, ITALY_PASSPORT_NUMBER, SPAIN_PASSPORT_NUMBER, UK_PASSPORT_NUMBER, USA_PASSPORT_NUMBER
Social Insurance Number (SIN)	CANADA_SOCIAL_INSURANCE_NUMBER
Social Security number (SSN)	SPAIN_SOCIAL_SECURITY_NUMBER, USA_SOCIAL_SECURITY_NUMBER
Taxpayer identification or reference number	AUSTRALIA_TAX_FILE_NUMBER, BRAZIL_CPF_NUMBER, FRANCE_TAX_IDENTIFICATION_NUMBER, GERMANY_TAX_IDENTIFICATION_NUMBER, SPAIN_NIE_NUMBER, SPAIN_NIF_NUMBER, SPAIN_TAX_IDENTIFICATION_NUMBER, USA_INDIVIDUAL_TAX_IDENTIFICATION_NUMBER

Updates to the default settings for automated sensitive data discovery

The following table describes changes to the settings that Amazon Macie uses by default for automated sensitive data discovery. For automatic alerts about these changes, subscribe to the RSS feed on the Macie document history page.

Change	Description	Date
Implemented a new, dynamic set of default managed data identifiers	New automated sensitive data discovery configurations are now based on a dynamic default set of managed data identifiers. If you enable automated sensitive data discovery for the first time on or after this date, your configuration is based on the dynamic set. If you enabled automated sensitive data discovery for the first time before this date, your configuration is based on a different set of managed data identifiers. For more information, see the notes after this table.	August 2, 2023
General availability	Initial release of automated sensitive data discovery.	November 28, 2022

If you initially enabled automated sensitive data discovery prior to August 2, 2023, your configuration isn't based on the dynamic set of default managed data identifiers. Instead, it's based on a static set of managed data identifiers that we defined for the initial release of automated sensitive data discovery, as listed in the table below.

To determine when you initially enabled automated sensitive data discovery you can use the Amazon Macie console: choose **Automated sensitive data discovery** in the navigation pane, and then refer to the enabled date in the **Status** section. You can also do this programmatically: use the GetAutomatedDiscoveryConfiguration operation of the Amazon Macie API and refer to the value for the `firstEnabledAt` field. If the date is prior to August 2, 2023, and you want to start using the dynamic set of default managed data identifiers, contact AWS Support for assistance.

The following table lists all the managed data identifiers that are in the static set. The table is sorted first by sensitive data category and then by sensitive data type. For details about specific managed data identifiers, see Using managed data identifiers.

Sensitive data category	Sensitive data type	Managed data identifier ID
Credentials	AWS secret access key	`AWS_CREDENTIALS`
Credentials	HTTP Basic Authorization header	`HTTP_BASIC_AUTH_HEADER`
Credentials	OpenSSH private key	`OPENSSH_PRIVATE_KEY`
Credentials	PGP private key	`PGP_PRIVATE_KEY`
Credentials	Public Key Cryptography Standard (PKCS) private key	`PKCS`
Credentials	PuTTY private key	`PUTTY_PRIVATE_KEY`
Financial information	Bank account number	`BANK_ACCOUNT_NUMBER` (for Canadian and US bank account numbers), `FRANCE_BANK_ACCOUNT_NUMBER`, `GERMANY_BANK_ACCOUNT_NUMBER`, `ITALY_BANK_ACCOUNT_NUMBER`, `SPAIN_BANK_ACCOUNT_NUMBER`, `UK_BANK_ACCOUNT_NUMBER`

Sensitive data category	Sensitive data type	Managed data identifier ID
Financial information	Credit card expiration date	`CREDIT_CARD_EXPIRATION`
Financial information	Credit card magnetic stripe data	`CREDIT_CARD_MAGNETIC_STRIPE`
Financial information	Credit card number	`CREDIT_CARD_NUMBER` (for credit card numbers in proximity of a keyword)
Financial information	Credit card verification code	`CREDIT_CARD_SECURITY_CODE`
Personal information: Personal health information (PHI)	Drug Enforcement Agency (DEA) Registration Number	`US_DRUG_ENFORCEMENT_AGENCY_NUMBER`
Personal information: PHI	Health Insurance Claim Number (HICN)	`USA_HEALTH_INSURANCE_CLAIM_NUMBER`
Personal information: PHI	Health insurance or medical identification number	`CANADA_HEALTH_NUMBER`, `EUROPEAN_HEALTH_INSURANCE_CARD_NUMBER`, `FINLAND_EUROPEAN_HEALTH_INSURANCE_NUMBER`, `FRANCE_HEALTH_INSURANCE_NUMBER`, `UK_NHS_NUMBER`, `USA_MEDICARE_BENEFICIARY_IDENTIFIER`
Personal information: PHI	Healthcare Common Procedure Coding System (HCPCS) code	`USA_HEALTHCARE_PROCEDURE_CODE`

Sensitive data category	Sensitive data type	Managed data identifier ID
Personal information: PHI	National Drug Code (NDC)	`USA_NATIONAL_DRUG_CODE`
Personal information: PHI	National Provider Identifier (NPI)	`USA_NATIONAL_PROVIDER_IDENTIFIER`
Personal information: PHI	Unique device identifier (UDI)	`MEDICAL_DEVICE_UDI`
Personal information: Personally identifiable information (PII)	Birth date	`DATE_OF_BIRTH`

Sensitive data category	Sensitive data type	Managed data identifier ID
Personal information: PII	Driver's license identification number	AUSTRALIA_DRIVERS_LICENSE, AUSTRIA_DRIVERS_LICENSE, BELGIUM_DRIVERS_LICENSE, BULGARIA_DRIVERS_LICENSE, CANADA_DRIVERS_LICENSE, CROATIA_DRIVERS_LICENSE, CYPRUS_DRIVERS_LICENSE, CZECHIA_DRIVERS_LICENSE, DENMARK_DRIVERS_LICENSE, DRIVERS_LICENSE (for the US), ESTONIA_DRIVERS_LICENSE, FINLAND_DRIVERS_LICENSE, FRANCE_DRIVERS_LICENSE, GERMANY_DRIVERS_LICENSE, GREECE_DRIVERS_LICENSE, HUNGARY_DRIVERS_LICENSE, IRELAND_DRIVERS_LICENSE, ITALY_DRIVERS_LICENSE, LATVIA_DRIVERS_LICENSE, LITHUANIA_DRIVERS_LICENSE, LUXEMBOURG_DRIVERS_LICENSE, MALTA_DRIVERS_LICENSE, NETHERLANDS_DRIVER

Sensitive data category	Sensitive data type	Managed data identifier ID
		S_LICENSE, POLAND_DRIVERS_LICENSE, PORTUGAL_DRIVERS_LICENSE, ROMANIA_DRIVERS_LICENSE, SLOVAKIA_DRIVERS_LICENSE, SLOVENIA_DRIVERS_LICENSE, SPAIN_DRIVERS_LICENSE, SWEDEN_DRIVERS_LICENSE, UK_DRIVERS_LICENSE
Personal information: PII	Electoral roll number	UK_ELECTORAL_ROLL_NUMBER
Personal information: PII	Full name	NAME
Personal information: PII	Global Positioning System (GPS) coordinates	LATITUDE_LONGITUDE
Personal information: PII	Mailing address	ADDRESS, BRAZIL_CEP_CODE
Personal information: PII	National identification number	BRAZIL_RG_NUMBER, FRANCE_NATIONAL_IDENTIFICATION_NUMBER, GERMANY_NATIONAL_IDENTIFICATION_NUMBER, ITALY_NATIONAL_IDENTIFICATION_NUMBER, SPAIN_DNI_NUMBER
Personal information: PII	National Insurance Number (NINO)	UK_NATIONAL_INSURANCE_NUMBER

Sensitive data category	Sensitive data type	Managed data identifier ID
Personal information: PII	Passport number	`CANADA_PASSPORT_NUMBER`, `FRANCE_PASSPORT_NUMBER`, `GERMANY_PASSPORT_NUMBER`, `ITALY_PASSPORT_NUMBER`, `SPAIN_PASSPORT_NUMBER`, `UK_PASSPORT_NUMBER`, `USA_PASSPORT_NUMBER`
Personal information: PII	Permanent residence number	`CANADA_NATIONAL_IDENTIFICATION_NUMBER`
Personal information: PII	Phone number	`BRAZIL_PHONE_NUMBER`, `FRANCE_PHONE_NUMBER`, `GERMANY_PHONE_NUMBER`, `ITALY_PHONE_NUMBER`, `PHONE_NUMBER` (for Canada and the US), `SPAIN_PHONE_NUMBER`, `UK_PHONE_NUMBER`
Personal information: PII	Social Insurance Number (SIN)	`CANADA_SOCIAL_INSURANCE_NUMBER`
Personal information: PII	Social Security number (SSN)	`SPAIN_SOCIAL_SECURITY_NUMBER`, `USA_SOCIAL_SECURITY_NUMBER`

Sensitive data category	Sensitive data type	Managed data identifier ID
Personal information: PII	Taxpayer identification or reference number	AUSTRALIA_TAX_FILE_NUMBER, BRAZIL_CNPJ_NUMBER, BRAZIL_CPF_NUMBER, FRANCE_TAX_IDENTIFICATION_NUMBER, GERMANY_TAX_IDENTIFICATION_NUMBER, SPAIN_NIE_NUMBER, SPAIN_NIF_NUMBER, SPAIN_TAX_IDENTIFICATION_NUMBER, UK_TAX_IDENTIFICATION_NUMBER, USA_INDIVIDUAL_TAX_IDENTIFICATION_NUMBER
Personal information: PII	Vehicle identification number (VIN)	VEHICLE_IDENTIFICATION_NUMBER

Running sensitive data discovery jobs

With Amazon Macie, you can create and run sensitive data discovery jobs to automate discovery, logging, and reporting of sensitive data in Amazon Simple Storage Service (Amazon S3) general purpose buckets. A *sensitive data discovery job* is a series of automated processing and analysis tasks that Macie performs to detect and report sensitive data in Amazon S3 objects. Each job provides detailed reports of the sensitive data that Macie finds and the analysis that Macie performs. By creating and running jobs, you can build and maintain a comprehensive view of the data that your organization stores in Amazon S3 and any security or compliance risks for that data.

To help you meet and maintain compliance with your data security and privacy requirements, Macie provides several options for scheduling and defining the scope of a job. You can configure a job to run only once for on-demand analysis and assessment, or on a recurring basis for periodic analysis, assessment, and monitoring. You also define the breadth and depth of a job's analysis—

specific S3 buckets that you select or buckets that match specific criteria. You can optionally refine the scope of that analysis by choosing additional options. The options include custom criteria that derive from properties of S3 objects, such as tags, prefixes, and when an object was last modified.

For each job, you also specify the types of sensitive data that you want Macie to detect and report. You can configure a job to use managed data identifiers that Macie provides, custom data identifiers that you define, or a combination of the two. By selecting specific managed and custom data identifiers for a job, you can tailor the analysis to focus on specific types of sensitive data. To fine tune the analysis, you can also configure a job to use allow lists. Allow lists specify text and text patterns that you want Macie to ignore, typically sensitive data exceptions for your organization's particular scenarios or environment.

Each job produces records of the sensitive data that Macie finds and the analysis that Macie performs—*sensitive data findings* and *sensitive data discovery results*. A *sensitive data finding* is a detailed report of sensitive data that Macie found in an S3 object. A *sensitive data discovery result* is a record that logs details about the analysis of an S3 object. Macie creates a sensitive data discovery result for each object that you configure a job to analyze. This includes objects that Macie doesn't find sensitive data in, and therefore don't produce sensitive data findings, and objects that Macie can't analyze due to errors or issues. Each type of record adheres to a standardized schema, which can help you query, monitor, and process the records to meet your security and compliance requirements.

Topics

Scope options for sensitive data discovery jobs

With sensitive data discovery jobs, you define the scope of the analysis that Amazon Macie performs to detect and report sensitive data in your Amazon Simple Storage Service (Amazon S3)

general purpose buckets. To help you do this, Macie provides several job-specific options that you can choose when you create and configure a job.

Scope options

- S3 buckets or bucket criteria
- Sampling depth
- Initial run: Include existing S3 objects
- S3 object criteria

S3 buckets or bucket criteria

When you create a sensitive data discovery job, you specify which S3 buckets store objects that you want Macie to analyze when the job runs. You can do this in two ways: by selecting specific S3 buckets from your bucket inventory, or by specifying custom criteria that derive from properties of S3 buckets.

Select specific S3 buckets

With this option, you explicitly select each S3 bucket to analyze. Then, when the job runs, Macie analyzes objects only in the buckets that you select. If you configure a job to run periodically on a daily, weekly, or monthly basis, Macie analyzes objects in those same buckets each time the job runs.

This configuration is helpful for cases where you want to perform targeted analysis of a specific set of data. It gives you precise, predictable control over which buckets a job analyzes.

Specify S3 bucket criteria

With this option, you define runtime criteria that determine which S3 buckets to analyze. The criteria consist of one or more conditions that derive from bucket properties, such as public access settings and tags. When the job runs, Macie identifies buckets that match your criteria, and then analyzes objects in those buckets. If you configure a job to run periodically, Macie does this each time the job runs. Consequently, Macie might analyze objects in different buckets each time the job runs, depending on changes to your bucket inventory and the criteria that you define.

This configuration is helpful for cases where you want the scope of the analysis to dynamically adapt to changes to your bucket inventory. If you configure a job to use bucket criteria and run

periodically, Macie automatically identifies new buckets that match the criteria and inspects those buckets for sensitive data.

The topics in this section provide additional details about each option.

Topics

Selecting specific S3 buckets

If you choose to explicitly select each S3 bucket that you want a job to analyze, Macie provides you with an inventory of your general purpose buckets in the current AWS Region. You can then review your inventory and select the buckets that you want. If you're the Macie administrator for an organization, your inventory includes buckets that your member accounts own. You can select as many as 1,000 of these buckets, spanning as many as 1,000 accounts.

To help you make your bucket selections, the inventory provides details and statistics for each bucket. This includes the amount of data that a job can analyze in each bucket—*classifiable objects* are objects that use a supported Amazon S3 storage class and have a file name extension for a supported file or storage format. The inventory also indicates whether you configured any existing jobs to analyze objects in a bucket. These details can help you estimate the breadth of a job and refine your bucket selections.

In the inventory table:

- **Sensitivity** – Specifies the bucket's current sensitivity score, if automated sensitive data discovery is enabled.
- **Classifiable objects** – Specifies the total number of objects that the job can analyze in the bucket.
- **Classifiable size** – Specifies the total storage size of all the objects that the job can analyze in the bucket.

  If the bucket stores compressed objects, this value doesn't reflect the actual size of those objects after they're decompressed. If versioning is enabled for the bucket, this value is based on the storage size of the latest version of each object in the bucket.

- **Monitored by job** – Specifies whether you configured any existing jobs to periodically analyze objects in the bucket on a daily, weekly, or monthly basis.

  If the value for this field is **Yes**, the bucket is explicitly included in a periodic job or the bucket matched the criteria for a periodic job within the past 24 hours. In addition, the status of at least one of those jobs is not *Cancelled*. Macie updates this data on a daily basis.
- **Latest job run** – If you configured any periodic or one-time jobs to analyze objects in the bucket, this field specifies the most recent date and time when one of those jobs started to run. Otherwise, a dash (–) appears in this field.

If the information icon () appears next to any bucket names, we recommend that you retrieve the latest bucket metadata from Amazon S3. To do this, choose refresh () above the table. The information icon indicates that a bucket was created during the past 24 hours, possibly after Macie last retrieved bucket and object metadata from Amazon S3 as part of the daily refresh cycle. For more information, see Data refreshes.

If the warning icon () appears next to a bucket's name, Macie isn't allowed to access the bucket or the bucket's objects. This means that the job won't be able to analyze objects in the bucket. To investigate the issue, review the bucket's policy and permissions settings in Amazon S3. For example, the bucket might have a restrictive bucket policy. For more information, see Allowing Macie to access S3 buckets and objects.

To customize your view and find specific buckets more easily, you can filter the table by entering filter criteria in the filter box. The following table provides some examples.

To show all buckets that...	Apply this filter...
Are owned by a specific account	**Account ID** = *the 12-digit ID for the account*
Are publicly accessible	**Effective permission** = **Public**

To show all buckets that...	Apply this filter...
Aren't included in any periodic jobs	**Actively monitored by job** = **False**
Aren't included in any periodic or one-time jobs	**Defined in job** = **False**
Have a specific tag key*	**Tag key** = *the tag key*
Have a specific tag value*	**Tag value** = *the tag value*
Store unencrypted objects (or objects that use client-side encryption)	**Object count by encryption** is **No encryption** and **From** = *1*

* Tag keys and values are case sensitive. Also, you have to specify a complete, valid value. You can't specify partial values or use wildcard characters.

To display additional details for a bucket, choose the bucket's name and refer to the details panel. In the panel, you can also:

- Pivot and drill down on certain fields by choosing a magnifying glass for the field. Choose  to show buckets with the same value. Choose  to show buckets with other values.
- Retrieve the latest metadata for objects in the bucket. This can be helpful if you recently created a bucket or made significant changes to the bucket's objects during the past 24 hours. To retrieve the data, choose refresh () in the **Object statistics** section of the panel. This option is available for buckets that store 30,000 or fewer objects.

In certain cases, the panel might not include all the details of a bucket. This can occur if you store more than 10,000 buckets in Amazon S3. Macie maintains complete inventory data for only 10,000 buckets for an account—the 10,000 buckets that were most recently created or changed. You can, however, configure a job to analyze objects in buckets that exceed this quota. To review additional details for these buckets, use Amazon S3.

Specifying S3 bucket criteria

If you choose to specify bucket criteria for a job, Macie provides options for defining and testing the criteria. These are runtime criteria that determine which S3 buckets store objects to analyze. Each time the job runs, Macie identifies general purpose buckets that match your criteria, and then analyzes objects in the appropriate buckets. If you're the Macie administrator for an organization, this includes buckets that your member accounts own.

Defining bucket criteria

Bucket criteria consist of one or more conditions that derive from properties of S3 buckets. Each condition, also referred to as a *criterion*, consists of the following parts:

- A property-based field, such as **Account ID** or **Effective permission**.
- An operator, either *equals* (eq) or *not equals* (neq).
- One or more values.
- An include or exclude statement that indicates whether to analyze (*include*) or skip (*exclude*) buckets that match the condition.

If you specify more than one value for a field, Macie uses OR logic to join the values. If you specify more than one condition for the criteria, Macie uses AND logic to join the conditions. In addition, exclude conditions take precedence over include conditions. For example, if you include buckets that are publicly accessible and exclude buckets that have specific tags, the job analyzes objects in any bucket that's publicly accessible unless the bucket has one of the specified tags.

You can define conditions that derive from any of the following property-based fields for S3 buckets.

Account ID

The unique identifier (ID) for the AWS account that owns a bucket. To specify multiple values for this field, enter the ID for each account and separate each entry with a comma.

Note that Macie doesn't support use of wildcard characters or partial values for this field.

Bucket name

The name of a bucket. This field correlates to the **Name** field, not the **Amazon Resource Name (ARN)** field, in Amazon S3. To specify multiple values for this field, enter the name of each bucket and separate each entry with a comma.

Note that values are case sensitive. In addition, Macie doesn't support use of wildcard characters or partial values for this field.

Effective permission

Specifies whether a bucket is publicly accessible. You can choose one or more of the following values for this field:

- **Not public** – The general public doesn't have read or write access to the bucket.
- **Public** – The general public has read or write access to the bucket.
- **Unknown** – Macie wasn't able to evaluate the public access settings for the bucket. An issue or quota prevented Macie from retrieving and evaluating the requisite data.

To determine whether a bucket is publicly accessible, Macie analyzes a combination of account- and bucket-level settings for the bucket: the block public access settings for the account; the block public access settings for the bucket; the bucket policy for the bucket; and, the access control list (ACL) for the bucket. For information about these settings, see Access control and Blocking public access to your Amazon S3 storage in the *Amazon Simple Storage Service User Guide*.

Shared access

Specifies whether a bucket is shared with another AWS account, an Amazon CloudFront origin access identity (OAI), or a CloudFront origin access control (OAC). You can choose one or more of the following values for this field:

- **External** – The bucket is shared with one or more of the following or any combination of the following: a CloudFront OAI, a CloudFront OAC, or an account that's external to (not part of) your organization.
- **Internal** – The bucket is shared with one or more accounts that are internal to (part of) your organization. It isn't shared with a CloudFront OAI or OAC.
- **Not shared** – The bucket isn't shared with another account, a CloudFront OAI, or a CloudFront OAC.
- **Unknown** – Macie wasn't able to evaluate the shared access settings for the bucket. An issue or quota prevented Macie from retrieving and evaluating the requisite data.

To determine whether a bucket is shared with another AWS account, Macie analyzes the bucket policy and ACL for the bucket. In addition, an *organization* is defined as a set of Macie accounts that are centrally managed as a group of related accounts through AWS Organizations or by

Macie invitation. For information about Amazon S3 options for sharing buckets, see Access control in the *Amazon Simple Storage Service User Guide*.

To determine whether a bucket is shared with a CloudFront OAI or OAC, Macie analyzes the bucket policy for the bucket. A CloudFront OAI or OAC allows users to access a bucket's objects through one or more specified CloudFront distributions. For information about CloudFront OAIs and OACs, see Restricting access to an Amazon S3 origin in the *Amazon CloudFront Developer Guide*.

Tags

The tags that are associated with a bucket. Tags are labels that you can define and assign to certain types of AWS resources, including S3 buckets. Each tag consists of a required tag key and an optional tag value. For information about tagging S3 buckets, see Using cost allocation S3 bucket tags in the *Amazon Simple Storage Service User Guide*.

For a sensitive data discovery job, you can use this type of condition to include or exclude buckets that have a specific tag key, a specific tag value, or a specific tag key and tag value (as a pair). For example:

- If you specify **`Project`** as a tag key and don't specify any tag values for a condition, any bucket that has the *Project* tag key matches the condition's criteria, regardless of the tag values that are associated with that tag key.
- If you specify **`Development`** and **`Test`** as tag values and don't specify any tag keys for a condition, any bucket that has the **`Development`** or **`Test`** tag value matches the condition's criteria, regardless of the tag keys that are associated with those tag values.

Tag keys and values are case sensitive. In addition, Macie doesn't support use of wildcard characters or partial values in tag conditions.

To specify multiple tag keys in a condition, enter each tag key in the **Key** field and separate each entry with a comma. To specify multiple tag values in a condition, enter each tag value in the **Value** field and separate each entry with a comma.

If you store more than 10,000 buckets in Amazon S3, note that Macie doesn't maintain tag data for all the buckets. Macie maintains complete inventory data for only 10,000 buckets for an account—the 10,000 buckets that were most recently created or changed. For all other buckets, any associated tag keys and values aren't included in inventory data. This means that the buckets won't match specific tag keys or values in a condition that uses the *equals* (eq) operator. If you specify a *not equals* (neq) operator for a tag-based condition, this means that the buckets will match the condition.

Testing bucket criteria

While you define your bucket criteria, you can test and refine the criteria by previewing the results. To do this, expand the **Preview the criteria results** section that appears below the criteria on the console. This section displays a table of up to 25 general purpose buckets that currently match the criteria.

The table also provides insight into the amount of data that the job can analyze in each bucket—*classifiable objects* are objects that use a supported Amazon S3 storage class and have a file name extension for a supported file or storage format. The table also indicates whether you configured any existing jobs to periodically analyze objects in a bucket.

In the table:

- **Sensitivity** – Specifies the bucket's current sensitivity score, if automated sensitive data discovery is enabled.
- **Classifiable objects** – Specifies the total number of objects that the job can analyze in the bucket.
- **Classifiable size** – Specifies the total storage size of all the objects that the job can analyze in the bucket.

  If the bucket stores compressed objects, this value doesn't reflect the actual size of those objects after they're decompressed. If versioning is enabled for the bucket, this value is based on the storage size of the latest version of each object in the bucket.
- **Monitored by job** – Specifies whether you configured any existing jobs to periodically analyze objects in the bucket on a daily, weekly, or monthly basis.

  If the value for this field is **Yes**, the bucket is explicitly included in a periodic job or the bucket matched the criteria for a periodic job within the past 24 hours. In addition, the status of at least one of those jobs is not *Cancelled*. Macie updates this data on a daily basis.

If the warning icon () appears next to a bucket's name, Macie isn't allowed to access the bucket or the bucket's objects. This means that the job won't be able to analyze objects in the bucket. To investigate the issue, review the bucket's policy and permissions settings in Amazon S3. For example, the bucket might have a restrictive bucket policy. For more information, see Allowing Macie to access S3 buckets and objects.

To refine the bucket criteria for the job, use the filter options to add, change, or remove conditions from the criteria. Macie then updates the table to reflect your changes.

Sampling depth

With this option, you specify the percentage of eligible S3 objects that you want a sensitive data discovery job to analyze. Eligible objects are objects that: use a supported Amazon S3 storage class, have a file name extension for a supported file or storage format, and match other criteria that you specify for the job.

If this value is less than 100%, Macie selects eligible objects to analyze at random, up to the specified percentage, and analyzes all the data in those objects. For example, if you configure a job to analyze 10,000 objects and you specify a sampling depth of 20%, Macie analyzes approximately 2,000 randomly selected, eligible objects when the job runs.

Reducing the sampling depth of a job can lower the cost and reduce the duration of a job. It's helpful for cases where the data in objects is highly consistent and you want to determine whether an S3 bucket, rather than each object, stores sensitive data.

Note that this option controls the percentage of *objects* that are analyzed, not the percentage of *bytes* that are analyzed. If you enter a sampling depth that's less than 100%, Macie analyzes all the data in each selected object, not that percentage of the data in each selected object.

Initial run: Include existing S3 objects

You can use sensitive data discovery jobs to perform ongoing, incremental analysis of objects in S3 buckets. If you configure a job to run periodically, Macie does this for you automatically—each run analyzes only those objects that were created or changed after the preceding run. With the **Include existing objects** option, you choose the starting point for the first increment:

- To analyze all existing objects immediately after you finish creating the job, select the checkbox for this option.
- To wait and analyze only those objects that are created or changed after you create the job and before the first run, clear the checkbox for this option.

  Clearing this checkbox is helpful for cases where you already analyzed the data and want to continue to analyze it periodically. For example, if you previously used another service or application to classify data and you recently started using Macie, you might use this option to ensure continued discovery and classification of your data without incurring unnecessary costs or duplicating classification data.

Each subsequent run of a periodic job automatically analyzes only those objects that are created or changed after the preceding run.

For both periodic and one-time jobs, you can also configure a job to analyze only those objects that are created or changed before or after a certain time or during a certain time range. To do this, add object criteria that use the last modified date for objects.

S3 object criteria

To fine tune the scope of a sensitive data discovery job, you can define custom criteria for S3 objects. Macie uses these criteria to determine which objects to analyze (*include*) or skip (*exclude*) when the job runs. The criteria consist of one or more conditions that derive from properties of S3 objects. The conditions apply to objects in all the S3 buckets that are included in the analysis. If a bucket stores multiple versions of an object, the conditions apply to the latest version of the object.

If you define multiple conditions as object criteria, Macie uses AND logic to join the conditions. In addition, exclude conditions take precedence over include conditions. For example, if you include objects that have the .pdf file name extension and exclude objects that are larger than 5 MB, the job analyzes any object that has the .pdf file name extension, unless the object is larger than 5 MB.

You can define conditions that derive from any of the following properties of S3 objects.

File name extension

This correlates to the file name extension of an S3 object. You can use this type of condition to include or exclude objects based on file type. To do this for multiple types of files, enter the file name extension for each type and separate each entry with a comma—for example: **`docx,pdf,xlsx`**. If you enter multiple file name extensions as values for a condition, Macie uses OR logic to join the values.

Note that values are case sensitive. In addition, Macie doesn't support the use of partial values or wildcard characters in this type of condition.

For information about the types of files that Macie can analyze, see Supported file and storage formats.

Last modified

This correlates to the **Last modified** field in Amazon S3. In Amazon S3, this field stores the date and time when an S3 object was created or last changed, whichever is latest.

For a sensitive data discovery job, this condition can be a specific date, a specific date and time, or an exclusive time range:

- To analyze objects that were last modified after a certain date or date and time, enter the values in the **From** fields.
- To analyze objects that were last modified before a certain date or date and time, enter the values in the **To** fields.
- To analyze objects that were last modified during a certain time range, use the **From** fields to enter the values for the first date or date and time in the time range. Use the **To** fields to enter the values for the last date or date and time in the time range.
- To analyze objects that were last modified at any time during a certain single day, enter the date in the **From** date field. Enter the date for the next day in the **To** date field. Then verify that both time fields are blank. (Macie treats a blank time field as `00:00:00`.) For example, to analyze objects that changed on August 9, 2023, enter **`2023/08/09`** in the **From** date field, enter **`2023/08/10`** in the **To** date field, and don't enter a value in either time field.

Enter any time values in Coordinated Universal Time (UTC) and use 24-hour notation.

Prefix

This correlates to the **Key** field in Amazon S3. In Amazon S3, this field stores the name of an S3 object, including the object's prefix. A *prefix* is similar to a directory path within a bucket. It enables you to group similar objects together in a bucket, much like you might store similar files together in a folder on a file system. For information about object prefixes and folders in Amazon S3, see Organizing objects in the Amazon S3 console using folders in the *Amazon Simple Storage Service User Guide*.

You can use this type of condition to include or exclude objects whose keys (names) begin with a certain value. For example, to exclude all objects whose key begins with *AWSLogs*, enter **`AWSLogs`** as the value for a **Prefix** condition, and then choose **Exclude**.

If you enter multiple prefixes as values for a condition, Macie uses OR logic to join the values. For example, if you enter **`AWSLogs1`** and **`AWSLogs2`** as values for a condition, any object whose key begins with *AWSLogs1* or *AWSLogs2* matches the condition's criteria.

When you enter a value for a **Prefix** condition, keep the following in mind:

- Values are case sensitive.
- Macie doesn't support the use of wildcard characters in these values.

- In Amazon S3, an object's key doesn't include the name of the bucket that stores the object. For this reason, don't specify bucket names in these values.
- If a prefix includes a delimiter, include the delimiter in the value. For example, enter **`AWSLogs/eventlogs`** to define a condition for all objects whose key begins with *AWSLogs/eventlogs*. Macie supports the default Amazon S3 delimiter, which is a slash (/), and custom delimiters.

Also note that an object matches a condition's criteria only if the object's key exactly matches the value that you enter, starting with the first character in the object's key. In addition, Macie applies a condition to the complete **Key** value for an object, including the object's file name.

For example, if an object's key is *AWSLogs/eventlogs/testlog.csv* and you enter any of the following values for a condition, the object matches the condition's criteria:

- **`AWSLogs`**
- **`AWSLogs/event`**
- **`AWSLogs/eventlogs/`**
- **`AWSLogs/eventlogs/testlog`**
- **`AWSLogs/eventlogs/testlog.csv`**

However, if you enter **`eventlogs`**, the object doesn't match the criteria—the condition's value doesn't include the first part of the key, *AWSLogs/*. Similarly, if you enter **`awslogs`**, the object doesn't match the criteria due to differences in capitalization.

Storage size

This correlates to the **Size** field in Amazon S3. In Amazon S3, this field indicates the total storage size of an S3 object. If an object is a compressed file, this value doesn't reflect the actual size of the file after it's decompressed.

You can use this type of condition to include or exclude objects that are smaller than a certain size, larger than a certain size, or fall within a certain size range. Macie applies this type of condition to all types of objects, including compressed or archive files and the files that they contain. For information about size-based restrictions for each supported format, see Quotas for Macie.

Tags

The tags that are associated with an S3 object. Tags are labels that you can define and assign to certain types of AWS resources, including S3 objects. Each tag consists of a required tag key and

an optional tag value. For information about tagging S3 objects, see Categorizing your storage using tags in the *Amazon Simple Storage Service User Guide*.

For a sensitive data discovery job, you can use this type of condition to include or exclude objects that have a specific tag. This can be a specific tag key or a specific tag key and tag value (as a pair). If you specify multiple tags as values for a condition, Macie uses OR logic to join the values. For example, if you specify **`Project1`** and **`Project2`** as tag keys for a condition, any object that has the *Project1* or *Project2* tag key matches the condition's criteria.

Note that tag keys and values are case sensitive. In addition, Macie doesn't support use of partial values or wildcard characters in this type of condition.

Creating a sensitive data discovery job

With Amazon Macie, you can create and run sensitive data discovery jobs to automate discovery, logging, and reporting of sensitive data in Amazon Simple Storage Service (Amazon S3) general purpose buckets. A *sensitive data discovery job* is a series of automated processing and analysis tasks that Macie performs to detect and report sensitive data in Amazon S3 objects. As the analysis progresses, Macie provides detailed reports of the sensitive data that it finds and the analysis that it performs: *sensitive data findings*, which report sensitive data that Macie finds in individual S3 objects, and *sensitive data discovery results*, which log details about the analysis of individual S3 objects. For more information, see Reviewing job results.

When you create a job, you start by specifying which S3 buckets store objects that you want Macie to analyze when the job runs—specific buckets that you select or buckets that match specific criteria. Then you specify how often to run the job—once, or periodically on a daily, weekly, or monthly basis. You can also choose options to refine the scope of the job's analysis. The options include custom criteria that derive from properties of S3 objects, such as tags, prefixes, and when an object was last modified.

After you define the schedule and scope of the job, you specify which managed data identifiers and custom data identifiers to use:

- A *managed data identifier* is a set of built-in criteria and techniques that are designed to detect a specific type of sensitive data—for example, credit card numbers, AWS secret access keys, or passport numbers for a particular country or region. These identifiers can detect a large and growing list of sensitive data types for many countries and regions, including multiple types of

credentials data, financial information, and personally identifiable information (PII). For more information, see Using managed data identifiers.

- A *custom data identifier* is a set of criteria that you define to detect sensitive data. With custom data identifiers, you can detect sensitive data that reflects your organization's particular scenarios, intellectual property, or proprietary data—for example, employee IDs, customer account numbers, or internal data classifications. You can supplement the managed data identifiers that Macie provides. For more information, see Building custom data identifiers.

You then optionally select allow lists to use. In Macie, an *allow list* specifies text or a text pattern to ignore. These are typically sensitive data exceptions for your particular scenarios or environment—for example, public names or phone numbers for your organization, or sample data that your organization uses for testing. For more information, see Defining sensitive data exceptions with allow lists.

When you finish choosing these options, you're ready to enter general settings for the job, such as the job's name and description. You can then review and save the job.

Tasks

- Before you begin: Set up key resources
- Step 1: Choose S3 buckets
- Step 2: Review your S3 bucket selections or criteria
- Step 3: Define the schedule and refine the scope
- Step 4: Select managed data identifiers
- Step 5: Select custom data identifiers
- Step 6: Select allow lists
- Step 7: Enter general settings
- Step 8: Review and create

Before you begin: Set up key resources

Before you create a job, it's a good idea to take the following steps:

- Verify that you configured a repository for your sensitive data discovery results. To do this, choose **Discovery results** in the navigation pane on the Amazon Macie console. To learn about these settings, see Storing and retaining sensitive data discovery results.

- Create any custom data identifiers that you want the job to use. To learn how, see Building custom data identifiers.
- Create any allow lists that you want the job to use. To learn how, see Defining sensitive data exceptions with allow lists.
- If you want to analyze S3 objects that are encrypted, ensure that Macie can access and use the appropriate encryption keys. For more information, see Analyzing encrypted S3 objects.
- If you want to analyze objects in an S3 bucket that has a restrictive bucket policy, ensure that Macie is allowed to access the objects. For more information, see Allowing Macie to access S3 buckets and objects.

If you do these things before you create a job, you streamline creation of the job and help ensure that the job can analyze the data that you want.

Step 1: Choose S3 buckets

When you create a job, the first step is to specify which S3 buckets store objects that you want Macie to analyze when the job runs. For this step, you have two options:

- **Select specific buckets** – With this option, you explicitly select each S3 bucket to analyze. Then, when the job runs, Macie analyzes objects only in the buckets that you select.
- **Specify bucket criteria** – With this option, you define runtime criteria that determine which S3 buckets to analyze. The criteria consist of one or more conditions that derive from bucket properties. Then, when the job runs, Macie identifies buckets that match your criteria and analyzes objects in those buckets.

For detailed information about these options, see Scope options for jobs.

The following sections provide instructions for choosing and configuring each option. Choose the section for the option that you want.

Select specific buckets

If you choose to explicitly select each S3 bucket to analyze, Macie provides you with an inventory of your general purpose buckets in the current AWS Region. You can then use this inventory to select one or more buckets for the job. To learn about this inventory, see Selecting specific S3 buckets.

If you're the Macie administrator for an organization, the inventory includes buckets that are owned by member accounts in your organization. You can select as many as 1,000 of these buckets, spanning as many as 1,000 accounts.

To select specific S3 buckets for the job

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Jobs**.
3. Choose **Create job**.
4. On the **Choose S3 buckets** page, choose **Select specific buckets**. Macie displays a table of all the general purpose buckets for your account in the current Region.
5. In the **Select S3 buckets** section, optionally choose refresh () to retrieve the latest bucket metadata from Amazon S3.

   If the information icon () appears next to any bucket names, we recommend that you do this. This icon indicates that a bucket was created during the past 24 hours, possibly after Macie last retrieved bucket and object metadata from Amazon S3 as part of the daily refresh cycle.
6. In the table, select the checkbox for each bucket that you want the job to analyze.

> **Tip**
>
> - To find specific buckets more easily, enter filter criteria in the filter box above the table. You can also sort the table by choosing a column heading.
> - To determine whether you already configured a job to periodically analyze objects in a bucket, refer to the **Monitored by job** field. If **Yes** appears in a field, the bucket is explicitly included in a periodic job or the bucket matched the criteria for a periodic job within the past 24 hours. In addition, the status of at least one of those jobs is not *Cancelled*. Macie updates this data on a daily basis.
> - To determine when an existing periodic or one-time job most recently analyzed objects in a bucket, refer to the **Latest job run** field. For additional information about that job, refer to the bucket's details.

- To display a bucket's details, choose the bucket's name. In addition to job-related information, the details panel provides statistics and other information about the bucket, such as the bucket's public access settings. To learn more about this data, see Reviewing your S3 bucket inventory.

7. When you finish selecting buckets, choose **Next**.

In the next step, you'll review and verify your selections.

Specify bucket criteria

If you choose to specify runtime criteria that determine which S3 buckets to analyze, Macie provides options to help you choose fields, operators, and values for individual conditions in the criteria. To learn more about these options, see Specifying S3 bucket criteria.

To specify S3 bucket criteria for the job

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Jobs**.
3. Choose **Create job**.
4. On the **Choose S3 buckets** page, choose **Specify bucket criteria**.
5. Under **Specify bucket criteria**, do the following to add a condition to the criteria:

   a. Place your cursor in the filter box, and then choose the bucket property to use for the condition.

   b. In the first box, choose an operator for the condition, **Equals** or **Not equals**.

   c. In the next box, enter one or more values for the property.

      Depending on the type and nature of the bucket property, Macie displays different options for entering values. For example, if you choose the **Effective permission** property, Macie displays a list of values to choose from. If you choose the **Account ID** property, Macie displays a text box in which you can enter one or more AWS account IDs. To enter multiple values in a text box, enter each value and separate each entry with a comma.

   d. Choose **Apply**. Macie adds the condition and displays it below the filter box.

      By default, Macie adds the condition with an include statement. This means that the job is configured to analyze (*include*) objects in buckets that match the condition. To skip

(*exclude*) buckets that match the condition, choose **Include** for the condition, and then choose **Exclude**.

e. Repeat the preceding steps for each additional condition that you want to add to the criteria.

6. To test your criteria, expand the **Preview the criteria results** section. This section displays a table of up to 25 general purpose buckets that currently match the criteria.
7. To refine your criteria, do any of the following:

   - To remove a condition, choose **X** for the condition.
   - To change a condition, remove the condition by choosing **X** for the condition. Then add a condition that has the correct settings.
   - To remove all conditions, choose **Clear filters**.

   Macie updates the table of criteria results to reflect your changes.
8. When you finish specifying bucket criteria, choose **Next**.

In the next step, you'll review and verify your criteria.

Step 2: Review your S3 bucket selections or criteria

For this step, verify that you chose the correct settings in the preceding step:

- **Review your bucket selections** - If you selected specific S3 buckets for the job, review the table of buckets and change your bucket selections as necessary. The table provides insight into the projected scope and cost of the job's analysis. The data is based on the size and types of objects that are currently stored in a bucket.

  In the table, the **Estimated cost** field indicates the total estimated cost (in US dollars) of analyzing objects in an S3 bucket. Each estimate reflects the projected amount of uncompressed data that the job will analyze in a bucket. If any objects are compressed or archive files, the estimate assumes that the files use a 3:1 compression ratio and the job can analyze all extracted files. For more information, see Forecasting and monitoring job costs.
- **Review your bucket criteria** - If you specified bucket criteria for the job, review each condition in the criteria. To change the criteria, choose **Previous**, and then use the filter options in the preceding step to enter the correct criteria. When you finish, choose **Next**.

When you finish reviewing and verifying the settings, choose **Next**.

Step 3: Define the schedule and refine the scope

For this step, specify how often you want the job to run—once, or periodically on a daily, weekly, or monthly basis. Also choose various options to refine the scope of the job's analysis. To learn about these options, see Scope options for jobs.

To define the schedule and refine the scope of the job

1. On the **Refine the scope** page, specify how often you want the job to run:
   - To run the job only once, immediately after you finish creating it, choose **One-time job**.
   - To run the job periodically on a recurring basis, choose **Scheduled job**. For **Update frequency**, choose whether to run the job daily, weekly, or monthly. Then use the **Include existing objects** option to define the scope of the job's first run:
     - Select this checkbox to analyze all existing objects immediately after you finish creating the job. Each subsequent run analyzes only those objects that are created or changed after the preceding run.
     - Clear this checkbox to skip analysis of all existing objects. The job's first run analyzes only those objects that are created or changed after you finish creating the job and before the first run starts. Each subsequent run analyzes only those objects that are created or changed after the preceding run.

       Clearing this checkbox is helpful for cases where you already analyzed the data and want to continue to analyze it periodically. For example, if you previously used another service or application to classify data and you recently started using Macie, you might use this option to ensure continued discovery and classification of your data without incurring unnecessary costs or duplicating classification data.
2. (Optional) To specify the percentage of objects that you want the job to analyze, enter the percentage in the **Sampling depth** box.

   If this value is less than 100%, Macie selects the objects to analyze at random, up to the specified percentage, and analyzes all the data in those objects. The default value is 100%.
3. (Optional) To add specific criteria that determine which S3 objects are included or excluded from the job's analysis, expand the **Additional settings** section, and then enter the criteria. These criteria consist of individual conditions that derive from properties of objects:

- To analyze (*include*) objects that meet a specific condition, enter the condition type and value, and then choose **Include**.
- To skip (*exclude*) objects that meet a specific condition, enter the condition type and value, and then choose **Exclude**.

Repeat this step for each include or exclude condition that you want.

If you enter multiple conditions, any exclude conditions take precedence over include conditions. For example, if you include objects that have the .pdf file name extension and exclude objects that are larger than 5 MB, the job analyzes any object that has the .pdf file name extension, unless the object is larger than 5 MB.

4. When you finish, choose **Next**.

Step 4: Select managed data identifiers

For this step, specify which managed data identifiers you want the job to use when it analyzes S3 objects. You have two options:

- **Use recommended settings** - With this option, the job analyzes S3 objects by using the set of managed data identifiers that we recommend for jobs. This set is designed to detect common categories and types of sensitive data. To review a list of managed data identifiers that are currently in the set, see Managed data identifiers recommended for jobs. We update that list each time we add or remove a managed data identifier from the set.
- **Use custom settings** - With this option, the job analyzes S3 objects by using managed data identifiers that you select. This can be all or only some of the managed data identifiers that are currently available. You can also configure the job to not use any managed data identifiers. The job can instead use only custom data identifiers that you select in the next step. To review a list of managed data identifiers that are currently available, see Quick reference: Managed data identifiers by type. We update that list each time we release a new managed data identifier.

When you choose either option, Macie displays a table of managed data identifiers. In the table, the **Sensitive data type** field specifies the unique identifier (ID) for a managed data identifier. This ID describes the type of sensitive data that the managed data identifier is designed to detect, for example: **USA_PASSPORT_NUMBER** for US passport numbers, **CREDIT_CARD_NUMBER** for

credit card numbers, and **PGP_PRIVATE_KEY** for PGP private keys. To find specific identifiers more quickly, you can sort and filter the table by sensitive data category or type.

To select managed data identifiers for the job

1. On the **Select managed data identifiers** page, under **Managed data identifier options**, do one of the following:

   - To use the set of managed data identifiers that we recommend for jobs, choose **Recommended**.

     If you choose this option and you configured the job to run more than once, each run automatically uses all the managed data identifiers that are in the recommended set when the run starts. This includes new managed data identifiers that we release and add to the set. It excludes managed data identifiers that we remove from the set and no longer recommend for jobs.

   - To use only specific managed data identifiers that you select, choose **Custom**, and then choose **Use specific managed data identifiers**. Then, in the table, select the checkbox for each managed data identifier that you want the job to use.

     If you choose this option and you configured the job to run more than once, each run uses only the managed data identifiers that you select. In other words, the job uses these same managed data identifiers each time it runs.

   - To use all the managed data identifiers that Macie currently provides, choose **Custom**, and then choose **Use specific managed data identifiers**. Then, in the table, select the checkbox in the selection column heading to select all rows.

     If you choose this option and you configured the job to run more than once, each run uses only the managed data identifiers that you select. In other words, the job uses these same managed data identifiers each time it runs.

   - To not use any managed data identifiers and use only custom data identifiers, choose **Custom**, and then choose **Don't use any managed data identifiers**. Then, in the next step, select the custom data identifiers to use.

2. When you finish, choose **Next**.

Step 5: Select custom data identifiers

For this step, select any custom data identifiers that you want the job to use when it analyzes S3 objects. The job will use the selected identifiers in addition to any managed data identifiers that you configured the job to use. To learn more about custom data identifiers, see Building custom data identifiers.

To select custom data identifiers for the job

1. On the **Select custom data identifiers** page, select the checkbox for each custom data identifier that you want the job to use. You can select as many as 30 custom data identifiers.

   > **Tip**
   >
   > To review or test the settings for a custom data identifier before you select it, choose the link icon () next to the identifier's name. Macie opens a page that displays the identifier's settings. You can also use this page to test the identifier with sample data. To do this, enter up to 1,000 characters of text in the **Sample data** box, and then choose **Test**. Macie evaluates the sample data by using the identifier, and then reports the number of matches.

2. When you finish selecting custom data identifiers, choose **Next**.

Step 6: Select allow lists

For this step, select any allow lists that you want the job to use when it analyzes S3 objects. To learn more about allow lists, see Defining sensitive data exceptions with allow lists.

To select allow lists for the job

1. On the **Select allow lists** page, select the checkbox for each allow list that you want the job to use. You can select as many as 10 lists.

> **Tip**
>
> To review the settings for an allow list before you select it, choose the link icon ( ) next to the list's name. Macie opens a page that displays the list's settings.
> If the list specifies a regular expression (*regex*), you can also use this page to test the regex with sample data. To do this, enter up to 1,000 characters of text in the **Sample data** box, and then choose **Test**. Macie evaluates the sample data by using the regex, and then reports the number of matches.

2. When you finish selecting allow lists, choose **Next**.

Step 7: Enter general settings

For this step, specify a name and, optionally, a description of the job. You can also assign tags to the job. A *tag* is a label that you define and assign to certain types of AWS resources. Each tag consists of a required tag key and an optional tag value. Tags can help you identify, categorize, and manage resources in different ways, such as by purpose, owner, environment, or other criteria. To learn more, see Tagging Macie resources.

To enter general settings for the job

1. On the **Enter general settings** page, enter a name for the job in the **Job name** box. The name can contain as many as 500 characters.
2. (Optional) For **Job description**, enter a brief description of the job. The description can contain as many as 200 characters.
3. (Optional) For **Tags**, choose **Add tag**, and then enter as many as 50 tags to assign to the job.
4. When you finish, choose **Next**.

Step 8: Review and create

For this final step, review the job's configuration settings and verify that they're correct. This is an important step. After you create a job, you can't change any of these settings. This helps ensure that you have an immutable history of sensitive data findings and discovery results for data privacy and protection audits or investigations that you perform.

Depending on the job's settings, you can also review the total estimated cost (in US dollars) of running the job once. If you selected specific S3 buckets for the job, the estimate is based on the size and types of objects in the buckets that you selected, and how much of that data the job can analyze. If you specified bucket criteria for the job, the estimate is based on the size and types of objects in as many as 500 buckets that currently match the criteria, and how much of that data the job can analyze. To learn about this estimate, see Forecasting and monitoring job costs.

To review and create the job

1. On the **Review and create** page, review each setting and verify that it's correct. To change a setting, choose **Edit** in the section that contains the setting, and then enter the correct setting. You can also use the navigation tabs to go to the page that contains a setting.
2. When you finish verifying the settings, choose **Submit** to create and save the job. Macie checks the settings and notifies you of any issues to address.

   > **Note**
   >
   > If you haven't configured a repository for your sensitive data discovery results, Macie displays a warning and doesn't save the job. To address this issue, choose **Configure** in the **Repository for sensitive data discovery results** section. Then enter the configuration settings for the repository. To learn how, see Storing and retaining sensitive data discovery results. After you enter the settings, return to the **Review and create** page and choose refresh ( ) in the **Repository for sensitive data discovery results** section of the page. Although we don't recommend it, you can temporarily override the repository requirement and save the job. If you do this, you risk losing discovery results from the job—Macie retains the results for only 90 days. To temporarily override the requirement, select the checkbox for the override option.

3. If Macie notifies you of issues to address, address the issues, and then choose **Submit** again to create and save the job.

If you configured the job to run once, on a daily basis, or on the current day of the week or month, Macie starts running the job immediately after you save it. Otherwise, Macie prepares to run the job on the specified day of the week or month. To monitor the job, you can check the status of the job.

Reviewing the results of a sensitive data discovery job

When you run a sensitive data discovery job, Amazon Macie automatically calculates and reports certain statistical data for the job. For example, Macie reports the number of times that the job has run, and the approximate number of Amazon Simple Storage Service (Amazon S3) objects that the job has yet to process during its current run. Macie also produces several types of results for the job: *log events*, *sensitive data findings*, and *sensitive data discovery results*.

Topics

Types of results for sensitive data discovery jobs

As a sensitive data discovery job progresses, Amazon Macie produces the following types of results for the job.

Log event

This is a record of an event that occurred while the job was running. Macie automatically logs and publishes data for certain events to Amazon CloudWatch Logs. The data in these logs provides a record of changes to the job's progress or status, such as the exact date and time when the job started or stopped running. The data also provides details about any account- or bucket-level errors that occurred while the job ran.

Log events can help you monitor a job and address any issues that prevented the job from analyzing the data that you want. If a job uses runtime criteria to determine which S3 buckets to analyze, log events can also help you determine whether and which S3 buckets matched the criteria when the job ran.

You can access log events by using the Amazon CloudWatch console or the Amazon CloudWatch Logs API. To help you navigate to the log events for a job, the Amazon Macie console provides a link to them. For more information, see Monitoring jobs with CloudWatch Logs.

Sensitive data finding

This is a report of sensitive data that Macie found in an S3 object. Each finding provides a severity rating and details such as:

- The date and time when Macie found the sensitive data.
- The category and types of sensitive data that Macie found.
- The number of occurrences of each type of sensitive data that Macie found.
- The unique identifier for the job that produced the finding.
- The name, public access settings, encryption type, and other information about the affected S3 bucket and object.

Depending on the affected S3 object's file type or storage format, the details can also include the location of as many as 15 occurrences of the sensitive data that Macie found. To report location data, sensitive data findings use a standardized JSON schema.

A sensitive data finding doesn't include the sensitive data that Macie found. Instead, it provides information that you can use for further investigation and remediation as necessary.

Macie stores sensitive data findings for 90 days. You can access them by using the Amazon Macie console or the Amazon Macie API. You can also monitor and process them by using other applications, services, and systems. For more information, see Reviewing and analyzing findings.

Sensitive data discovery result

This is a record that logs details about the analysis of an S3 object. Macie automatically creates a sensitive data discovery result for each object that you configure a job to analyze. This includes objects that Macie doesn't find sensitive data in, and therefore don't produce sensitive data findings, and objects that Macie can't analyze due to errors or issues such as permissions settings or use of an unsupported file or storage format.

If Macie finds sensitive data in an S3 object, the sensitive data discovery result includes data from the corresponding sensitive data finding. It provides additional information too, such as the location of as many as 1,000 occurrences of each type of sensitive data that Macie found in the object. For example:

- The column and row number for a cell or field in a Microsoft Excel workbook, CSV file, or TSV file
- The path to a field or array in a JSON or JSON Lines file
- The line number for a line in a non-binary text file other than a CSV, JSON, JSON Lines, or TSV file—for example, an HTML, TXT, or XML file
- The page number for a page in an Adobe Portable Document Format (PDF) file

- The record index and the path to a field in a record in an Apache Avro object container or Apache Parquet file

If the affected S3 object is an archive file, such as a .tar or .zip file, the sensitive data discovery result also provides detailed location data for occurrences of sensitive data in individual files that Macie extracted from the archive. Macie doesn't include this information in sensitive data findings for archive files. To report location data, sensitive data discovery results use a standardized JSON schema.

A sensitive data discovery result doesn't include the sensitive data that Macie found. Instead, it provides you with an analysis record that can be helpful for data privacy and protection audits or investigations.

Macie stores your sensitive data discovery results for 90 days. You can't access them directly on the Amazon Macie console or with the Amazon Macie API. Instead, you configure Macie to encrypt and store them in an S3 bucket. The bucket can serve as a definitive, long-term repository for all of your sensitive data discovery results. You can then optionally access and query the results in that repository. To learn how to configure these settings, see Storing and retaining sensitive data discovery results.

After you configure the settings, Macie writes your sensitive data discovery results to JSON Lines (.jsonl) files, and it encrypts and adds those files to the S3 bucket as GNU Zip (.gz) files. To help you navigate to the results, the Amazon Macie console provides links to them.

Sensitive data findings and sensitive data discovery results both adhere to standardized schemas. This can help you optionally query, monitor, and process them by using other applications, services, and systems.

Tips

For a detailed, instructional example of how you might query and use sensitive data discovery results to analyze and report potential data security risks, see the following blog post on the *AWS Security Blog*: How to query and visualize Macie sensitive data discovery results with Amazon Athena and Amazon QuickSight.

For samples of Amazon Athena queries that you can use to analyze sensitive data discovery results, visit the Amazon Macie Results Analytics repository on GitHub. This repository also provides instructions for configuring Athena to retrieve and decrypt your results, and scripts for creating tables for the results.

Reviewing statistics and results for a sensitive data discovery job

To review processing statistics and the results of a sensitive data discovery job, you can use the Amazon Macie console or the Amazon Macie API. Follow these steps to review the statistics and results by using the console.

To access a job's processing statistics programmatically, use the DescribeClassificationJob operation of the Amazon Macie API. For programmatic access to the findings that a job produced, use the ListFindings operation and specify the job's unique identifier in a filter condition for the `classificationDetails.jobId` field. To learn how, see Creating and applying filters to Macie findings. You can then use the GetFindings operation to retrieve the details of the findings.

To review statistics and results for a job

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Jobs**.
3. On the **Jobs** page, choose the name of the job whose statistics and results you want to review. The details panel displays statistics, settings, and other information about the job.
4. In the details panel, do any of the following:

   - To review processing statistics for the job, refer to the **Statistics** section of the panel. This section displays statistics such as the number of times that the job has run, and the approximate number of objects that the job has yet to process during its current run.
   - To review log events for the job, choose **Show results** at the top of the panel, and then choose **Show CloudWatch logs**. Macie opens the Amazon CloudWatch console and displays a table of the log events that Macie published for the job.
   - To review all the sensitive data findings that the job produced, choose **Show results** at the top of the panel, and then choose **Show findings**. Macie opens the **Findings** page and displays all the findings from the job. To review the details of a particular finding, choose the finding, and then refer to the details panel.

     > **Tip**
     >
     > In the finding details panel, you can use the link in the **Detailed result location** field to navigate to the corresponding sensitive data discovery result in Amazon S3:

- If the finding applies to a large archive or compressed file, the link displays the folder that contains the discovery results for the file. An archive or compressed file is *large* if it generates more than 100 discovery results.
- If the finding applies to a small archive or compressed file, the link displays the file that contains the discovery results for the file. An archive or compressed file is *small* if it generates 100 or fewer discovery results.
- If the finding applies to another type of file, the link displays the file that contains the discovery results for the file.

- To review all the sensitive data discovery results that the job produced, choose **Show results** at the top of the panel, and then choose **Show classifications**. Macie opens the Amazon S3 console and displays the folder that contains all the discovery results for the job. This option is available only after you configure Macie to store your sensitive data discovery results in an S3 bucket.

Managing sensitive data discovery jobs

To help you manage your sensitive data discovery jobs, Amazon Macie maintains a complete inventory of your jobs in each AWS Region. With this inventory, you can manage your jobs as a single collection, and access configuration settings, processing statistics, and the status of individual jobs.

For example, you can identify all the jobs that you configured to run on a recurring basis for periodic analysis, assessment, and monitoring. You can also review a breakdown of the configuration settings for a job. This includes settings that define the scope of the analysis. It also includes settings that specify the types of sensitive data that you want Macie to detect and report when the job runs. If you use the Amazon Macie console to manage your jobs, each job's details also provide direct access to sensitive data findings and other results that the job produced.

In addition to these tasks, you can create custom variations of individual jobs. You can copy an existing job, adjust the settings for the copy, and then save the copy as a new job. This can be helpful for cases where you want to analyze different sets of data in the same way, or the same set of data in different ways. It can also be helpful if you want to adjust the configuration settings for an existing job—cancel the existing job, copy it, and then adjust and save the copy as a new job.

Topics

- Reviewing the settings for a sensitive data discovery job
- Checking the status of a sensitive data discovery job
- Changing the status of a sensitive data discovery job
- Copying a sensitive data discovery job

Reviewing your inventory of sensitive data discovery jobs

On the Amazon Macie console, you can review a complete inventory of your sensitive data discovery jobs in the current AWS Region. The inventory provides both summary information for all of your jobs and details about individual jobs. Summary information includes: the current status of each job; whether a job runs on a scheduled, periodic basis; and, whether a job is configured to analyze objects in specific Amazon Simple Storage Service (Amazon S3) buckets or S3 buckets that match runtime criteria. For individual jobs, you can also access details such as a breakdown of the job's configuration settings. If a job has already run, the details also provide direct access to sensitive data findings and other types of results that the job produced.

To review your job inventory

Follow these steps to review your job inventory by using the Amazon Macie console. To access your inventory programmatically, use the ListClassificationJobs operation of the Amazon Macie API.

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Jobs**. The **Jobs** page opens and displays the number of jobs in your inventory and a table of those jobs.
3. At the top of the page, optionally choose refresh () to retrieve the current status of each job.
4. In the **Jobs** table, review summary information for your jobs:
   - **Job name** – The name of the job.
   - **Resources** – Whether the job is configured to analyze objects in specific S3 buckets or buckets that match runtime criteria. If you explicitly selected buckets for the job to analyze, this field indicates the number of buckets that you selected. If you configured the job to use runtime criteria, the value for this field is **Criteria based**.
   - **Job type** – Whether the job is configured to run once (**One time**) or on a scheduled, periodic basis (**Scheduled**).

- **Status** – The current status of the job. To learn more about this value, see Checking the status of a job.
- **Created at** – When the job was created.

5. To analyze your inventory or find a specific job more quickly, do any of the following:

   - To sort the table by a specific field, choose the column heading for the field. To change the sort order, choose the column heading again.
   - To show only those jobs that have a specific value for a field, place your cursor in the filter box. In the menu that appears, choose the field to use for the filter, and enter the value for the filter. Then choose **Apply**.
   - To hide jobs that have a specific value for a field, place your cursor in the filter box. In the menu that appears, choose the field to use for the filter, and enter the value for the filter. Then choose **Apply**. In the filter box, choose the equals icon (●) for the filter. This changes the filter's operator from *equals* to *not equals* (⦸).
   - To remove a filter, choose the remove filter icon (⊗) for the filter to remove.

6. To review additional settings and details for a particular job, choose the job's name. Then refer to the details panel. For information about these details, see Reviewing configuration settings for a job.

Reviewing the settings for a sensitive data discovery job

On the Amazon Macie console, you can use the details panel on the **Jobs** page to review configuration settings and other information about individual sensitive data discovery jobs. For example, you can review a list of the Amazon Simple Storage Service (Amazon S3) buckets that a job is configured to analyze. You can also determine which managed and custom data identifiers a job is configured to use when analyzing objects in those buckets.

Note that you can't change any configuration settings for an existing job. This helps ensure that you have an immutable history of sensitive data findings and discovery results for data privacy and protection audits or investigations that you perform.

If you want to change an existing job, you can cancel the job. Then copy the job, configure the copy to use the settings that you want, and save the copy as a new job. If you do this, you should also take steps to ensure that the new job doesn't analyze existing data in the same way again. To do this, note the date and time when you cancel the existing job. Then configure the scope of the new job to include only those objects that are created or changed after you cancel the original job. For example, you can use object criteria to define an exclude condition that specifies when you cancelled the original job.

To review the configuration settings for a job

Follow these steps to review a job's configuration settings by using the Amazon Macie console. To review the settings programmatically, use the DescribeClassificationJob operation of the Amazon Macie API.

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Jobs**. The **Jobs** page opens and displays the number of jobs in your inventory and a table of those jobs.
3. In the **Jobs** table, choose the name of the job whose settings you want to review. To find the job more quickly, you can filter the table by using the filter options above the table. You can also sort the table in ascending or descending order by certain fields.

When you choose a job in the table, the details panel displays the job's configuration settings and other information about the job. Depending on the job's settings, the panel contains the following sections.

General information

This section provides general information about the job. For example, it shows the Amazon Resource Name (ARN) of the job, when the job most recently started to run, and the current status of the job. If you paused the job, this section also indicates when you paused the job, and when the job or latest job run expired or will expire if you don't resume it.

Statistics

This section shows processing statistics for the job. For example, it specifies the number of times that the job has run, and the approximate number of S3 objects that the job has yet to process during its current run.

Scope

This section indicates how often the job runs. It also shows settings that refine the job's scope—for example, the sampling depth, and any object criteria that include or exclude S3 objects from the analysis.

S3 buckets

This section appears in the panel if the job is configured to analyze buckets that you explicitly selected when you created the job. It indicates the number of AWS accounts that the job is configured to analyze data for. It also indicates the number of buckets that the job is configured to analyze and the names of those buckets (grouped by account).

To show the complete list of accounts and buckets in JSON format, choose the number in the **Total buckets** field.

S3 bucket criteria

This section appears in the panel if the job uses runtime criteria to determine which buckets to analyze. It lists the criteria that the job is configured to use. To show the criteria in JSON format, choose **Details**. Then choose the **Criteria** tab in the window that appears.

To review a list of buckets that currently match the criteria, choose **Details**. Then choose the **Matching buckets** tab in the window that appears. Optionally choose refresh ( ) to retrieve the latest data. The tab lists up to 25 buckets that currently match the criteria.

> **Tip**
>
> If the job has already run, you can also determine whether any buckets matched the criteria when the job ran and, if so, the names of those buckets. To do this, review log events for the job: choose **Show results** at the top of the panel, and then choose **Show CloudWatch logs**. Macie opens the Amazon CloudWatch console and displays a table of log events for the job. The events include a `BUCKET_MATCHED_THE_CRITERIA` event for each bucket that matched the criteria and was included in the job's analysis. For more information, see Monitoring jobs with CloudWatch Logs.

Custom data identifiers

This section appears in the panel if the job is configured to use one or more custom data identifiers. It specifies the names of those custom data identifiers.

Allow lists

This section appears in the panel if the job is configured to use one or more allow lists. It specifies the names of those lists. To review the settings and status of a list, choose the link icon ( ) next to the list's name.

Managed data identifiers

This section indicates which managed data identifiers the job is configured to use. This is determined by the managed data identifier selection type for the job:

- **Recommended** – Use the managed data identifiers that are in the recommended set when the job runs.
- **Include selected** – Use only the managed data identifiers listed in the **Selections** section.
- **Include all** – Use all the managed data identifiers that are available when the job runs.
- **Exclude selected** – Use all the managed data identifiers that are available when the job runs, except the ones listed in the **Selections** section.
- **Exclude all** – Don't use any managed data identifiers. Use only the specified custom data identifiers.

To review these settings in JSON format, choose **Details**.

Tags

This section appears in the panel if tags are assigned to the job. It lists those tags. A *tag* is a label that you define and assign to certain types of AWS resources. Each tag consists of a required tag key and an optional tag value. To learn more, see Tagging Macie resources.

To review and save the job's settings in JSON format, choose the unique identifier for the job (**Job ID**) at the top of the panel. Then choose **Download**.

Checking the status of a sensitive data discovery job

When you create a sensitive data discovery job, its initial status is **Active (Running)** or **Active (Idle)**, depending on the job's type and schedule. The job then passes through additional states, which you can monitor as the job progresses.

> **Tip**
>
> In addition to monitoring the overall status of a job, you can monitor specific events that occur as a job progresses. You can do this by using logging data that Amazon Macie automatically publishes to Amazon CloudWatch Logs. The data in these logs provides a record of changes to a job's status and details about any account- or bucket-level errors that occur while a job runs. For more information, see Monitoring jobs with CloudWatch Logs.

To check the status of a job

Follow these steps to check the status of a job by using the Amazon Macie console. To check a job's status programmatically, use the DescribeClassificationJob operation of the Amazon Macie API.

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Jobs**. The **Jobs** page opens and displays the number of jobs in your inventory and a table of those jobs.
3. At the top of the page, choose refresh () to retrieve the current status of each job.
4. In the **Jobs** table, locate the job whose status you want to check. To find the job more quickly, you can filter the table by using the filter options above the table. You can also sort the table in ascending or descending order by certain fields.
5. Refer to the **Status** field in the table. This field indicates the job's current status.

A job's status can be one of the following.

Active (Idle)

For a periodic job, the previous run is complete and the next scheduled run is pending. This value doesn't apply to one-time jobs.

Active (Running)

For a one-time job, the job is currently in progress. For a periodic job, a scheduled run is in progress.

Cancelled

For any type of job, the job was stopped permanently (cancelled).

A job has this status if you explicitly cancelled it or, if it's a one-time job, you paused the job and didn't resume it within 30 days. A job can also have this status if you previously suspended Macie in the current AWS Region.

Complete

For a one-time job, the job ran successfully and is now complete. This value doesn't apply to periodic jobs. Instead, the status of a periodic job changes to **Active (Idle)** when each run completes successfully.

Paused (By Macie)

For any type of job, the job was stopped temporarily (paused) by Macie.

A job has this status if completion of the job or a job run would exceed the monthly sensitive data discovery quota for your account. When this happens, Macie automatically pauses the job. Macie automatically resumes the job when the next calendar month starts and the monthly quota is reset for your account, or you increase the quota for your account.

If you're the Macie administrator for an organization and you configured the job to analyze data for member accounts, the job can also have this status if completion of the job or a job run would exceed the monthly sensitive data discovery quota for a member account.

If a job is running and the analysis of eligible objects reaches this quota for a member account, the job stops analyzing objects that are owned by the account. When the job finishes analyzing objects for all other accounts that haven't met the quota, Macie automatically pauses the job. If it's a one-time job, Macie automatically resumes the job when the next calendar month starts or the quota is increased for all the affected accounts, whichever occurs first. If it's a periodic job, Macie automatically resumes the job when the next run is scheduled to start or the next calendar month starts, whichever occurs first. If a scheduled run starts before the next calendar month starts or the quota is increased for an affected account, the job doesn't analyze objects that are owned by the account.

Paused (By user)

For any type of job, the job was stopped temporarily (paused) by you.

If you pause a one-time job and you don't resume it within 30 days, the job expires and Macie cancels it. If you pause a periodic job while it's actively running and you don't resume it within

30 days, the job's run expires and Macie cancels the run. To check the expiration date for a paused job or job run, choose the job's name in the table, and then refer to the **Expires** field in the **Status details** section of the details panel.

If a job is cancelled or paused, you can refer to the job's details to determine whether the job started to run or, for a periodic job, ran at least once before it was cancelled or paused. To do this, choose the job's name in the **Jobs** table, and then refer to the details panel. In the panel, the **Number of runs** field indicates the number of times that the job has run. The **Last run time** field indicates the most recent date and time when the job started to run.

Depending on the job's current status, you can optionally pause, resume, or cancel the job. For more information, see Changing the status of a job.

Changing the status of a sensitive data discovery job

After you create a sensitive data discovery job, you can pause it temporarily or cancel it permanently. When you pause a job that's actively running, Amazon Macie immediately begins to pause all processing tasks for the job. When you cancel a job that's actively running, Macie immediately begins to stop all processing tasks for the job. You can't resume or restart a job after it's cancelled.

If you pause a one-time job, you can resume it within 30 days. When you resume the job, Macie immediately resumes processing from the point where you paused the job. Macie doesn't restart the job from the beginning. If you don't resume a one-time job within 30 days of pausing it, the job expires and Macie cancels it.

If you pause a periodic job, you can resume it at any time. If you resume a periodic job and the job was idle when you paused it, Macie resumes the job according to the schedule and other configuration settings that you chose when you created the job. If you resume a periodic job and the job was actively running when you paused it, how Macie resumes the job depends on when you resume the job:

- If you resume the job within 30 days of pausing it, Macie immediately resumes the latest scheduled run from the point where you paused the job. Macie doesn't restart the run from the beginning.
- If you don't resume the job within 30 days of pausing it, the latest scheduled run expires and Macie cancels all remaining processing tasks for the run. When you subsequently resume the job,

Macie resumes the job according to the schedule and other configuration settings that you chose when you created the job.

To help you determine when a paused job or job run will expire, Macie adds an expiration date to the job's details while the job is paused. In addition, we notify you approximately seven days before the job or job run will expire. We notify you again when the job or job run expires and is cancelled. To notify you, we send email to the address that's associated with your AWS account. We also create AWS Health events and Amazon CloudWatch Events for your account. To check the expiration date by using the console, choose the job's name in the table on the **Jobs** page. Then refer to the **Expires** field in the **Status details** section of the details panel. To check the date programmatically, use the DescribeClassificationJob operation of the Amazon Macie API.

To pause, resume, or cancel a job

To pause, resume, or cancel a job by using the Amazon Macie console, follow these steps. To do this programmatically, use the UpdateClassificationJob operation of the Amazon Macie API.

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Jobs**. The **Jobs** page opens and displays the number of jobs in your inventory and a table of those jobs.
3. At the top of the page, choose refresh
   ()
   to retrieve the current status of each job.
4. In the **Jobs** table, select the checkbox for the job that you want to pause, resume, or cancel. To find the job more quickly, you can filter the table by using the filter options above the table. You can also sort the table in ascending or descending order by certain fields.
5. On the **Actions** menu, do one of the following:

   - To pause the job temporarily, choose **Pause**. This option is available only if the job's current status is **Active (Idle)**, **Active (Running)**, or **Paused (By Macie)**.
   - To resume the job, choose **Resume**. This option is available only if the job's current status is **Paused (By user)**.
   - To cancel the job permanently, choose **Cancel**. If you choose this option, you can't subsequently resume or restart the job.

Copying a sensitive data discovery job

To quickly create a sensitive data discovery job that's similar to an existing job, you can create a copy of the existing job. You can then edit the copy's settings, and save the copy as a new job. This can be helpful for cases where you want to analyze different sets of data in the same way, or the same set of data in different ways. It can also be helpful if you want to adjust the configuration settings for an existing job—cancel the existing job, copy it, and then adjust and save the copy as a new job.

To copy a job

Follow these steps to copy a job by using the Amazon Macie console. To copy a job programmatically, use the DescribeClassificationJob operation of the Amazon Macie API to retrieve the configuration settings for the job that you want to copy. Then use the CreateClassificationJob operation to create a copy of the job.

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Jobs**. The **Jobs** page opens and displays the number of jobs in your inventory and a table of those jobs.
3. In the **Jobs** table, select the checkbox for the job that you want to copy. To find the job more quickly, you can filter the table by using the filter options above the table. You can also sort the table in ascending or descending order by certain fields.
4. On the **Actions** menu, choose **Copy to new**.
5. Complete the steps on the console to review and adjust the settings for the copy of the job. For the **Refine the scope** step, consider choosing options that prevent the job from analyzing existing data in the same way again:

   - For a one-time job, use object criteria to include only those objects that were created or changed after a certain time. For example, if you're creating a copy of a job that you cancelled, add a **Last modified** condition that specifies the date and time when you cancelled the existing job.
   - For a periodic job, clear the **Include existing objects** checkbox. If you do this, the first run of the job analyzes only those objects that are created or changed after you create the job and before the job's first run. You can also use object criteria to exclude objects that were last modified before a certain date and time.

   For additional details about this and other steps, see Creating a sensitive data discovery job.

6. When you finish, choose **Submit** to save the copy as a new job.

If you configured the job to run once, on a daily basis, or on the current day of the week or month, Macie starts running the job immediately after you save it. Otherwise, Macie prepares to run the job on the specified day of the week or month. To monitor the job, you can check the status of the job.

Monitoring sensitive data discovery jobs with CloudWatch Logs

In addition to monitoring the overall status of a sensitive data discovery job, you can monitor and analyze specific events that occur as a job progresses. You can do this by using near real-time logging data that Amazon Macie automatically publishes to Amazon CloudWatch Logs. The data in these logs provides a record of changes to a job's progress or status. For example, you can use the data to determine the exact date and time when a job started to run, was paused, or finished running.

The log data also provides details about any account- or bucket-level errors that occur while a job runs. For example, Macie logs an event if the permissions settings for an Amazon Simple Storage Service (Amazon S3) bucket prevent a job from analyzing objects in the bucket. The event indicates when the error occurred, and it identifies the affected bucket and the AWS account that owns the bucket. The data for these types of events can help you identify, investigate, and address errors that prevent Macie from analyzing the data that you want.

With Amazon CloudWatch Logs, you can monitor, store, and access log files from multiple systems, applications, and AWS services, including Macie. You can also query and analyze log data, and configure CloudWatch Logs to notify you when certain events occur or thresholds are met. CloudWatch Logs also provides features for archiving log data and exporting the data to Amazon S3. To learn more about CloudWatch Logs, see the Amazon CloudWatch Logs User Guide.

Topics

- How logging works for sensitive data discovery jobs
- Reviewing logs for sensitive data discovery jobs
- Understanding log events for sensitive data discovery jobs

How logging works for sensitive data discovery jobs

When you start running sensitive data discovery jobs, Amazon Macie automatically creates and configures the appropriate resources in Amazon CloudWatch Logs to log events for all of your jobs. Macie then publishes event data to those resources automatically when your jobs run. The permissions policy for the Macie service-linked role for your account allows Macie to perform these tasks on your behalf. You don't need to take any steps to create or configure resources in CloudWatch Logs to log event data for your jobs.

In CloudWatch Logs, logs are organized into *log groups*. Each log group contains *log streams*. Each log stream contains *log events*. The general purpose of each of these resources is as follows:

- A *log group* is a collection of log streams that share the same retention, monitoring, and access control settings—for example, the collection of logs for all of your sensitive data discovery jobs.
- A *log stream* is a sequence of log events that share the same source—for example, an individual sensitive data discovery job.
- A *log event* is a record of an activity that was recorded by an application or resource—for example, an individual event that Macie recorded and published for a particular sensitive data discovery job.

Macie publishes events for all of your sensitive data discovery jobs to one log group. Each job has a unique log stream in that log group. The log group has the following prefix and name:

```
/aws/macie/classificationjobs
```

If this log group already exists, Macie uses it to store log events for your jobs. This can be helpful if your organization uses automated configuration, such as AWS CloudFormation, to create log groups with predefined retention periods, encryption settings, tags, metric filters, and so on, for job events.

If this log group doesn't exist, Macie creates it with the default settings that CloudWatch Logs uses for new log groups. The settings include a log retention period of **Never Expire**, which means that CloudWatch Logs stores the logs indefinitely. You can change the retention period for the log group. To learn how, see Working with log groups and log streams in the *Amazon CloudWatch Logs User Guide*.

Within this log group, Macie creates a unique log stream for each job that you run, the first time that the job runs. The name of the log stream is the unique identifier for the job, such as `85a55dc0fa6ed0be5939d0408example`, in the following format:

```
/aws/macie/classificationjobs/85a55dc0fa6ed0be5939d0408example
```

Each log stream contains all the log events that Macie recorded and published for the corresponding job. For periodic jobs, this includes events for all of the job's runs. If you delete the log stream for a periodic job, Macie creates the stream again the next time that the job runs. If you delete the log stream for a one-time job, you can't restore it.

Note that logging is enabled by default for all of your jobs. You can't disable it or otherwise prevent Macie from publishing job events to CloudWatch Logs. If you don't want to store the logs, you can reduce the retention period for the log group to as little as one day. At the end of the retention period, CloudWatch Logs automatically deletes expired event data from the log group.

Reviewing logs for sensitive data discovery jobs

After you start running sensitive data discovery jobs in Amazon Macie, you can review logs for your jobs by using Amazon CloudWatch Logs. CloudWatch Logs provides features that are designed to help you review, analyze, and monitor log data. You can use these features to work with log streams and events for jobs as you would work with any other type of log data in CloudWatch Logs.

For example, you can search and filter aggregate data to identify specific types of events that occurred for all of your jobs during a specific time range. Or you can perform a targeted review of all the events that occurred for a particular job. CloudWatch Logs also provides options for monitoring log data, defining metric filters, and creating custom alarms.

> **Tip**
>
> To quickly navigate to the log data for a particular job, you can use the Amazon Macie console. To do this, choose the job's name on the **Jobs** page. At the top of the details panel, choose **Show results**, and then choose **Show CloudWatch logs**. Macie opens the Amazon CloudWatch console and displays a table of log events for the job.

To review logs for sensitive data discovery jobs

Follow these steps to navigate to and review log data by using the Amazon CloudWatch console. To review the data programmatically, use the Amazon CloudWatch Logs API.

1. Open the CloudWatch console at https://console.aws.amazon.com/cloudwatch/.

2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you ran jobs that you want to review logs for.
3. In the navigation pane, choose **Logs**, and then choose **Log groups**.
4. On the **Log groups** page, choose the **/aws/macie/classificationjobs** log group. CloudWatch displays a table of log streams for the jobs that you've run. There is one unique stream for each job. The name of each stream correlates to the unique identifier for a job.
5. On the **Log streams** tab, do one of the following:

   - To review the log events for a particular job, choose the log stream for the job. To find the stream more easily, enter the job's unique identifier in the filter box above the table. After you choose the log stream, CloudWatch displays a table of log events for the job.
   - To review log events for all of your jobs, choose **Search all log streams**. CloudWatch displays a table of log events for all of your jobs.
6. (Optional) In the filter box above the table, enter terms, phrases, or values that specify characteristics of specific events to review. For more information, see Search log data using filter patterns in the *Amazon CloudWatch Logs User Guide*.
7. To review the details of a specific log event, choose expand (▶) in the row for the event. CloudWatch displays the event's details in JSON format. To learn more about these details, see Understanding log events for jobs.

As you familiarize yourself with the data in the log events, you can perform additional tasks to streamline analysis and monitoring of the data. For example, you can create metrics filters that turn log data into numerical CloudWatch metrics. You can also create custom alarms that make it easier to identify and respond to specific log events. For more information, see the Amazon CloudWatch Logs User Guide.

Understanding log events for sensitive data discovery jobs

To help you monitor your sensitive data discovery jobs, Amazon Macie automatically publishes logging data for jobs to Amazon CloudWatch Logs. The data in these logs provides a record of changes to a job's progress or status. For example, you can use the data to determine the exact date and time when a job started to run or finished running. The data also provides details about certain types of errors that can occur while a job runs. This data can help you identify, investigate, and address errors that prevent Macie from analyzing the data that you want.

When you start running jobs, Macie automatically creates and configures the appropriate resources in CloudWatch Logs to log events for all of your jobs. Macie then publishes event data to those resources automatically when your jobs run. For more information, see How logging works for jobs.

By using CloudWatch Logs, you can then query and analyze log data for your jobs. For example, you can search and filter aggregate data to identify specific types of events that occurred for all of your jobs during a specific time range. Or you can perform a targeted review of all the events that occurred for a particular job. CloudWatch Logs also provides options for monitoring log data, defining metric filters, and creating custom alarms. For example, you can configure CloudWatch Logs to notify you if a certain type of event occurs when your jobs run. For more information, see the Amazon CloudWatch Logs User Guide.

Topics

Log event schema for sensitive data discovery jobs

Each log event for a sensitive data discovery job is a JSON object that contains a standard set of fields and conforms to the Amazon CloudWatch Logs event schema. Some types of events have additional fields that provide information that's particularly useful for that type of event. For example, events for account-level errors include the account ID for the affected AWS account. Events for bucket-level errors include the name of the affected Amazon Simple Storage Service (Amazon S3) bucket.

The following example shows the log event schema for sensitive data discovery jobs. In this example, the event reports that Amazon Macie wasn't able to analyze any objects in an S3 bucket because Amazon S3 denied access to the bucket.

```
{
    "adminAccountId": "123456789012",
    "jobId": "85a55dc0fa6ed0be5939d0408example",
    "eventType": "BUCKET_ACCESS_DENIED",
    "occurredAt": "2024-04-14T17:11:30.574809Z",
    "description": "Macie doesn’t have permission to access the affected S3 bucket.",
```

```
    "jobName": "My_Macie_Job",
    "operation": "ListObjectsV2",
    "runDate": "2024-04-14T17:08:30.345809Z",
    "affectedAccount": "111122223333",
    "affectedResource": {
        "type": "S3_BUCKET_NAME",
        "value": "amzn-s3-demo-bucket"
    }
}
```

In the preceding example, Macie attempted to list the bucket's objects by using the ListObjectsV2 operation of the Amazon S3 API. When Macie sent the request to Amazon S3, Amazon S3 denied access to the bucket.

The following fields are common to all log events for sensitive data discovery jobs:

- `adminAccountId` – The unique identifier for the AWS account that created the job.
- `jobId` – The unique identifier for the job.
- `eventType` – The type of event that occurred.
- `occurredAt` – The date and time, in Coordinated Universal Time (UTC) and extended ISO 8601 format, when the event occurred.
- `description` – A brief description of the event.
- `jobName` – The name of the job.

Depending on the type and nature of an event, a log event can also contain the following fields:

- `affectedAccount` – The unique identifier for the AWS account that owns the affected resource.
- `affectedResource` – A JSON object that provides details about the affected resource. In the object, the `type` field specifies a field that stores metadata about a resource. The `value` field specifies the value for the field (`type`).
- `operation` – The operation that Macie attempted to perform and caused the error.
- `runDate` – The date and time, in Coordinated Universal Time (UTC) and extended ISO 8601 format, when the applicable job or job run started.

Types of log events for sensitive data discovery jobs

Amazon Macie publishes log events for three categories of events that can occur for a sensitive data discovery job:

- Job status events, which record changes to the status or progress of a job or a job run.
- Account-level error events, which record errors that prevented Macie from analyzing Amazon S3 data for a specific AWS account.
- Bucket-level error events, which record errors that prevented Macie from analyzing data in a specific S3 bucket.

The topics in this section list and describe the types of events that Macie publishes for each category.

Job status events

A job status event records a change to the status or progress of a job or a job run. For periodic jobs, Macie logs and publishes these events for both the overall job and individual job runs.

The following example uses sample data to show the structure and nature of the fields in a job status event. In this example, a `SCHEDULED_RUN_COMPLETED` event indicates that a scheduled run of a periodic job finished running. The run started on April 14, 2024, at 17:09:30 UTC, as indicated by the `runDate` field. The run finished on April 14, 2024, at 17:16:30 UTC, as indicated by the `occurredAt` field.

```
{
    "adminAccountId": "123456789012",
    "jobId": "ffad0e71455f38a4c7c220f3cexample",
    "eventType": "SCHEDULED_RUN_COMPLETED",
    "occurredAt": "2024-04-14T17:16:30.574809Z",
    "description": "The scheduled job run finished running.",
    "jobName": "My_Daily_Macie_Job",
    "runDate": "2024-04-14T17:09:30.574809Z"
}
```

The following table lists and describes the types of job status events that Macie logs and publishes to CloudWatch Logs. The **Event type** column indicates the name of each event as it appears in the `eventType` field of an event. The **Description** column provides a brief description of the event as it appears in the `description` field of an event. The **Additional information** provides

information about the type of job that the event applies to. The table is sorted first by the general chronological order in which events might occur, and then in ascending alphabetical order by event type.

Event type	Description	Additional information
JOB_CREATED	The job was created.	Applies to one-time and periodic jobs.
ONE_TIME_JOB_STARTED	The job started running.	Applies only to one-time jobs.
SCHEDULED_RUN_STARTED	The scheduled job run started running.	Applies only to periodic jobs. To log the start of a one-time job, Macie publishes a ONE_TIME_JOB_STARTED event, not this type of event.
BUCKET_MATCHED_THE_CRITERIA	The affected bucket matched the bucket criteria specified for the job.	Applies to one-time and periodic jobs that use runtime bucket criteria to determine which S3 buckets to analyze. The `affectedResource` object specifies the name of the bucket that matched the criteria and was included in the job's analysis.
NO_BUCKETS_MATCHED_THE_CRITERIA	The job started running but no buckets currently match the bucket criteria specified for the job. The job didn't analyze any data.	Applies to one-time and periodic jobs that use runtime bucket criteria to determine which S3 buckets to analyze.

Event type	Description	Additional information
SCHEDULED_RUN_COMPLETED	The scheduled job run finished running.	Applies only to periodic jobs. To log completion of a one-time job, Macie publishes a JOB_COMPLETED event, not this type of event.
JOB_PAUSED_BY_USER	The job was paused by a user.	Applies to one-time and periodic jobs that you stopped temporarily (paused).
JOB_RESUMED_BY_USER	The job was resumed by a user.	Applies to one-time and periodic jobs that you stopped temporarily (paused) and later resumed.
JOB_PAUSED_BY_MACIE_SERVICE_QUOTA_MET	The job was paused by Macie. Completion of the job would exceed a monthly quota for the affected account.	Applies to one-time and periodic jobs that Macie stopped temporarily (paused). Macie automatically pauses a job when additional processing by the job or a job run would exceed the monthly sensitive data discovery quota for one or more accounts that the job analyzes data for. To avoid this issue, consider increasing the quota for the affected accounts.

Event type	Description	Additional information
JOB_RESUMED_BY_MACIE_SERVICE_QUOTA_LIFTED	The job was resumed by Macie. The monthly service quota was lifted for the affected account.	Applies to one-time and periodic jobs that Macie stopped temporarily (paused) and later resumed. If Macie automatically paused a one-time job, Macie automatically resumes the job when the subsequent month starts or the monthly sensitive data discovery quota is increased for all the affected accounts, whichever occurs first. If Macie automatically paused a periodic job, Macie automatically resumes the job when the next run is scheduled to start or the subsequent month starts, whichever occurs first.

Event type	Description	Additional information
JOB_CANCELLED	The job was cancelled.	Applies to one-time and periodic jobs that you stopped permanently (cancelled) or, for one-time jobs, paused and didn't resume within 30 days. If you suspend or disable Macie, this type of event also applies to jobs that were active or paused when you suspended or disabled Macie. Macie automatically cancels your jobs in an AWS Region if you suspend or disable Macie in the Region.
JOB_COMPLETED	The job finished running.	Applies only to one-time jobs. To log completion of a job run for a periodic job, Macie publishes a SCHEDULED_RUN_COMPLETED event, not this type of event.

Account-level error events

An account-level error event records an error that prevented Macie from analyzing objects in S3 buckets that are owned by a specific AWS account. The `affectedAccount` field in each event specifies the account ID for that account.

The following example uses sample data to show the structure and nature of the fields in an account-level error event. In this example, an `ACCOUNT_ACCESS_DENIED` event indicates that Macie wasn't able to analyze objects in any S3 buckets that are owned by account `444455556666`.

```
{
    "adminAccountId": "123456789012",
    "jobId": "85a55dc0fa6ed0be5939d0408example",
    "eventType": "ACCOUNT_ACCESS_DENIED",
    "occurredAt": "2024-04-14T17:08:30.585709Z",
    "description": "Macie doesn't have permission to access S3 bucket data for the
 affected account.",
    "jobName": "My_Macie_Job",
    "operation": "ListBuckets",
    "runDate": "2024-04-14T17:05:27.574809Z",
    "affectedAccount": "444455556666"
}
```

The following table lists and describes the types of account-level error events that Macie logs and publishes to CloudWatch Logs. The **Event type** column indicates the name of each event as it appears in the `eventType` field of an event. The **Description** column provides a brief description of the event as it appears in the `description` field of an event. The **Additional information** column provides any applicable tips for investigating or addressing the error that occurred. The table is sorted in ascending alphabetical order by event type.

Event type	Description	Additional information
ACCOUNT_ACCESS_DENIED	Macie doesn't have permission to access S3 bucket data for the affected account.	This typically occurs because the buckets that are owned by the account have restrictive bucket policies. For information about how to address this issue, see Allowing Macie to access S3 buckets and objects. The value for the `operation` field in the event can help you determine which permissions settings prevented Macie from accessing S3 data for the account. This field indicates

Event type	Description	Additional information
		the Amazon S3 operation that Macie attempted to perform when the error occurred.
ACCOUNT_DISABLED	The job skipped resources that are owned by the affected account. Macie was disabled for the account.	To address this issue, re-enable Macie for the account in the same AWS Region.
ACCOUNT_DISASSOCIATED	The job skipped resources that are owned by the affected account. The account isn't associated with your Macie administrator account as a member account anymore.	This occurs if you, as a Macie administrator for an organization, configure a job to analyze data for a member account and the account is later removed from your organization. To address this issue, re-associate the affected account with your Macie administrator account as a member account. For more information, see Managing multiple accounts.
ACCOUNT_ISOLATED	The job skipped resources that are owned by the affected account. The AWS account was isolated.	–

Event type	Description	Additional information
ACCOUNT_REGION_DISABLED	The job skipped resources that are owned by the affected account. The AWS account isn't active in the current AWS Region.	–
ACCOUNT_SUSPENDED	The job was cancelled or skipped resources that are owned by the affected account. Macie was suspended for the account.	If the specified account is your own account, Macie automatically cancelled the job when you suspended Macie in the same Region. To address the issue, re-enable Macie in the Region. If the specified account is a member account, re-enable Macie for that account in the same Region.
ACCOUNT_TERMINATED	The job skipped resources that are owned by the affected account. The AWS account was terminated.	–

Bucket-level error events

A bucket-level error event records an error that prevented Macie from analyzing objects in a specific S3 bucket. The `affectedAccount` field in each event specifies the account ID for the AWS account that owns the bucket. The `affectedResource` object in each event specifies the name of the bucket.

The following example uses sample data to show the structure and nature of the fields in a bucket-level error event. In this example, a `BUCKET_ACCESS_DENIED` event indicates that Macie wasn't able to analyze any objects in the S3 bucket named `amzn-s3-demo-bucket`. When Macie

attempted to list the bucket's objects by using the ListObjectsV2 operation of the Amazon S3 API, Amazon S3 denied access to the bucket.


```
{
    "adminAccountId": "123456789012",
    "jobId": "85a55dc0fa6ed0be5939d0408example",
    "eventType": "BUCKET_ACCESS_DENIED",
    "occurredAt": "2024-04-14T17:11:30.574809Z",
    "description": "Macie doesn't have permission to access the affected S3 bucket.",
    "jobName": "My_Macie_Job",
    "operation": "ListObjectsV2",
    "runDate": "2024-04-14T17:09:30.685209Z",
    "affectedAccount": "111122223333",
    "affectedResource": {
        "type": "S3_BUCKET_NAME",
        "value": "amzn-s3-demo-bucket"
    }
}
```


The following table lists and describes the types of bucket-level error events that Macie logs and publishes to CloudWatch Logs. The **Event type** column indicates the name of each event as it appears in the `eventType` field of an event. The **Description** column provides a brief description of the event as it appears in the `description` field of an event. The **Additional information** column provides any applicable tips for investigating or addressing the error that occurred. The table is sorted in ascending alphabetical order by event type.

Event type	Description	Additional information
BUCKET_ACCESS_DENIED	Macie doesn't have permission to access the affected S3 bucket.	This typically occurs because a bucket has a restrictive bucket policy. For information about how to address this issue, see Allowing Macie to access S3 buckets and objects. The value for the `operation` field in the event can help you determine which permissions settings

Event type	Description	Additional information
		prevented Macie from accessing the bucket. This field indicates the Amazon S3 operation that Macie attempted to perform when the error occurred.
BUCKET_DETAILS_UNAVAILABLE	A temporary issue prevented Macie from retrieving details about the bucket and the bucket's objects.	This occurs if a transient issue prevented Macie from retrieving the bucket and object metadata that it needs to analyze a bucket's objects. For example, an Amazon S3 exception occurred when Macie tried to verify that it's allowed to access the bucket. To address the issue for a one-time job, consider creating and running a new, one-time job to analyze objects in the bucket. For a scheduled job, Macie will try to retrieve the metadata again during the next job run.
BUCKET_DOES_NOT_EXIST	The affected S3 bucket doesn't exist anymore.	This typically occurs because a bucket was deleted.
BUCKET_IN_DIFFERENT_REGION	The affected S3 bucket was moved to a different AWS Region.	–

Event type	Description	Additional information
BUCKET_OWNER_CHANGED	The owner of the affected S3 bucket changed. Macie doesn't have permission to access the bucket anymore.	This typically occurs if ownership of a bucket was transferred to an AWS account that isn't part of your organization. The `affectedAccount` field in the event indicates the account ID for the account that previously owned the bucket.

Forecasting and monitoring costs for sensitive data discovery jobs

Amazon Macie pricing is based partly on the amount of data that you analyze by running sensitive data discovery jobs. To forecast and monitor your estimated costs for running sensitive data discovery jobs, you can review cost estimates that Macie provides when you create a job and after you start running jobs.

To review and monitor your actual costs, you can use AWS Billing and Cost Management. AWS Billing and Cost Management provides features that are designed to help you track and analyze your costs for AWS services, and manage budgets for your account or organization. It also provides features that can help you forecast usage costs based on historical data. To learn more, see the AWS Billing User Guide.

For information about Macie pricing, see Amazon Macie pricing.

Topics

- Forecasting the cost of a sensitive data discovery job
- Monitoring estimated costs for sensitive data discovery jobs

Forecasting the cost of a sensitive data discovery job

When you create a sensitive data discovery job, Amazon Macie can calculate and display estimated costs during two key steps in the job creation process: when you review the table of S3 buckets

that you selected for the job (step 2) and when you review all the settings for the job (step 8). These estimates can help you determine whether to adjust the job's settings before you save the job. The availability and nature of the estimates depends on the settings that you choose for the job.

Reviewing estimated costs for individual buckets (step 2)

If you explicitly select individual buckets for a job to analyze, you can review the estimated cost of analyzing objects in each of those buckets. Macie displays these estimates during step 2 of the job creation process, when you review your bucket selections. In the table for this step, the **Estimated cost** field indicates the total estimated cost (in US dollars) of running the job once to analyze objects in a bucket.

Each estimate reflects the projected amount of uncompressed data that the job will analyze in a bucket, based on the size and types of objects that are currently stored in the bucket. The estimate also reflects Macie pricing for the current AWS Region.

Only classifiable objects are included in the cost estimate for a bucket. A *classifiable object* is an S3 object that uses a supported Amazon S3 storage class and has a file name extension for a supported file or storage format. If any classifiable objects are compressed or archive files, the estimate assumes that the files use a 3:1 compression ratio and the job can analyze all extracted files.

Reviewing the total estimated cost of a job (step 8)

If you create a one-time job or you create and configure a periodic job to include existing S3 objects, Macie calculates and displays the job's total estimated cost during the final step of the job creation process. You can review this estimate while you review and verify all the settings that you selected for the job.

This estimate indicates the total projected cost (in US dollars) of running the job once in the current Region. The estimate reflects the projected amount of uncompressed data that the job will analyze. It's based on the size and types of objects that are currently stored in buckets that you explicitly selected for the job or up to 500 buckets that currently match bucket criteria that you specified for the job, depending on the job's settings.

Note that this estimate doesn't reflect any options that you selected to refine and reduce the scope of the job—for example, a lower sampling depth, or criteria that exclude certain S3 objects from the job. It also doesn't reflect your monthly sensitive data discovery quota, which

might limit the scope and cost of the job's analysis, or any discounts that might apply to your account.

In addition to the total estimated cost of the job, the estimate provides aggregated data that offers insight into the projected scope and cost of the job:

- **Size** values indicate the total storage size of the objects that the job can and can't analyze.
- **Object count** values indicate the total number of objects that the job can and can't analyze.

In these values, a **Classifiable** object is an S3 object that uses a supported Amazon S3 storage class and has a file name extension for a supported file or storage format. Only classifiable objects are included in the cost estimate. A **Not classifiable** object is an object that doesn't use a supported storage class or doesn't have a file name extension for a supported file or storage format. These objects aren't included in the cost estimate.

The estimate provides additional aggregated data for S3 objects that are compressed or archive files. The **Compressed** value indicates the total storage size of objects that use a supported Amazon S3 storage class and have a file name extension for a supported type of compressed or archive file. The **Uncompressed** value indicates the approximate size of these objects if they're decompressed, based on a specified compression ratio. This data is relevant due to the way that Macie analyzes compressed files and archive files.

When Macie analyzes a compressed or archive file, it inspects both the full file and the contents of the file. To inspect the file's contents, Macie decompresses the file, and then inspects each extracted file that uses a supported format. The actual amount of data that a job analyzes therefore depends on:

- Whether a file uses compression and, if so, the compression ratio that it uses.
- The number, size, and format of the extracted files.

By default, Macie assumes the following when it calculates cost estimates for a job:

- All compressed and archive files use a 3:1 compression ratio.
- All the extracted files use a supported file or storage format.

These assumptions can result in a larger size estimate for the scope of the data that the job will analyze, and, consequently, a higher cost estimate for the job.

You can recalculate the job's total estimated cost based on a different compression ratio. To do this, choose the ratio from the **Choose an estimated compression ratio** list in the **Estimated cost** section. Macie then updates the estimate to match your selection.

For more information about how Macie calculates estimated costs, see Understanding estimated usage costs.

Monitoring estimated costs for sensitive data discovery jobs

If you're already running sensitive data discovery jobs, the **Usage** page on the Amazon Macie console can help you monitor the estimated cost of those jobs. The page shows your estimated costs (in US dollars) for using Macie in the current AWS Region during the current calendar month. For information about how Macie calculates these estimates, see Understanding estimated usage costs.

To review your estimated costs for running jobs

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to review your estimated costs.
3. In the navigation pane, choose **Usage**.
4. On the **Usage** page, refer to the breakdown of estimated costs for your account. The **Sensitive data discovery jobs** item reports the total estimated cost of the jobs that you've run thus far during the current month in the current Region.

   If you're the Macie administrator for an organization, the **Estimated costs** section shows estimated costs for your organization overall for the current month in the current Region. To show the total estimated cost of the jobs that were run for a specific account, choose the account in the table. The **Estimated costs** section then shows a breakdown of estimated costs for the account, including the estimated cost of the jobs that were run. To show this data for a different account, choose the account in the table. To clear your account selection, choose **X** next to the account ID.

To review and monitor your actual costs, use AWS Billing and Cost Management.

Managed data identifiers recommended for sensitive data discovery jobs

To optimize the results of your sensitive data discovery jobs, you can configure individual jobs to automatically use the set of managed data identifiers that we recommend for jobs. A *managed data identifier* is a set of built-in criteria and techniques that are designed to detect a specific type

of sensitive data—for example, AWS secret access keys, credit card numbers, or passport numbers for a particular country or region.

The recommended set of managed data identifiers is designed to detect common categories and types of sensitive data. Based on our research, it can detect general categories and types of sensitive data while also optimizing your job results by reducing noise. As we release new managed data identifiers, we add them to this set if they're likely to further optimize your job results. Over time, we might also add or remove existing managed data identifiers from the set. If we add or remove a managed data identifier from the recommended set, we update this page to indicate the nature and timing of the change. For automatic alerts about these changes, you can subscribe to the RSS feed on the Macie document history page.

When you create a sensitive data discovery job, you specify which managed data identifiers you want the job to use to analyze objects in Amazon Simple Storage Service (Amazon S3) buckets. To configure a job to use the recommended set of managed data identifiers, choose the *Recommended* option when you create the job. The job will then automatically use all the managed data identifiers that are in the recommended set when the job starts to run. If you configure a job to run more than once, each run will automatically use all the managed data identifiers that are in the recommended set when the run starts.

The following topics list the managed data identifiers that are currently in the recommended set, organized by sensitive data category and type. They specify the unique identifier (ID) for each managed data identifier in the set. This ID describes the type of sensitive data that a managed data identifier is designed to detect, for example: `PGP_PRIVATE_KEY` for PGP private keys and `USA_PASSPORT_NUMBER` for US passport numbers.

Topics

For details about specific managed data identifiers or a complete list of all the managed data identifiers that Macie currently provides, see Using managed data identifiers.

Credentials

To detect occurrences of credentials data in S3 objects, the recommended set uses the following managed data identifiers.

Sensitive data type	Managed data identifier ID
AWS secret access key	`AWS_CREDENTIALS`
HTTP Basic Authorization header	`HTTP_BASIC_AUTH_HEADER`
OpenSSH private key	`OPENSSH_PRIVATE_KEY`
PGP private key	`PGP_PRIVATE_KEY`
Public Key Cryptography Standard (PKCS) private key	`PKCS`
PuTTY private key	`PUTTY_PRIVATE_KEY`

Financial information

To detect occurrences of financial information in S3 objects, the recommended set uses the following managed data identifiers.

Sensitive data type	Managed data identifier ID
Credit card magnetic stripe data	`CREDIT_CARD_MAGNETIC_STRIPE`
Credit card number	`CREDIT_CARD_NUMBER` (for credit card numbers in proximity of a keyword)

Personally identifiable information (PII)

To detect occurrences of personally identifiable information (PII) in S3 objects, the recommended set uses the following managed data identifiers.

Sensitive data type	Managed data identifier ID
Driver's license identification number	CANADA_DRIVERS_LICENSE, DRIVERS_LICENSE (for the US), UK_DRIVERS_LICENSE
Electoral roll number	UK_ELECTORAL_ROLL_NUMBER
National identification number	FRANCE_NATIONAL_IDENTIFICATION_NUMBER, GERMANY_NATIONAL_IDENTIFICATION_NUMBER, ITALY_NATIONAL_IDENTIFICATION_NUMBER, SPAIN_DNI_NUMBER
National Insurance Number (NINO)	UK_NATIONAL_INSURANCE_NUMBER
Passport number	CANADA_PASSPORT_NUMBER, FRANCE_PASSPORT_NUMBER, GERMANY_PASSPORT_NUMBER, ITALY_PASSPORT_NUMBER, SPAIN_PASSPORT_NUMBER, UK_PASSPORT_NUMBER, USA_PASSPORT_NUMBER
Social Insurance Number (SIN)	CANADA_SOCIAL_INSURANCE_NUMBER
Social Security number (SSN)	SPAIN_SOCIAL_SECURITY_NUMBER, USA_SOCIAL_SECURITY_NUMBER
Taxpayer identification or reference number	AUSTRALIA_TAX_FILE_NUMBER, BRAZIL_CPF_NUMBER, FRANCE_TAX_IDENTIFICATION_NUMBER, GERMANY_TAX_IDENTIFICATION_NUMBER, SPAIN_NIE_NUMBER, SPAIN_NIF_NUMBER, SPAIN_TAX_IDENTIFICATION_NUMBER, USA_INDIVIDUAL_TAX_IDENTIFICATION_NUMBER

Updates to the recommended set

The following table describes changes to the set of managed data identifiers that we recommend for sensitive data discovery jobs. For automatic alerts about these changes, subscribe to the RSS feed on the Macie document history page.

Change	Description	Date
General availability	Initial release of the recommended set.	June 27, 2023

Analyzing encrypted Amazon S3 objects

When you enable Amazon Macie for your AWS account, Macie creates a service-linked role that grants Macie the permissions that it requires to call Amazon Simple Storage Service (Amazon S3) and other AWS services on your behalf. A service-linked role simplifies the process of setting up an AWS service because you don't have to manually add permissions for the service to complete actions on your behalf. To learn about this type of role, see IAM roles in the *AWS Identity and Access Management User Guide*.

The permissions policy for the Macie service-linked role (`AWSServiceRoleForAmazonMacie`) allows Macie to perform actions that include retrieving information about your S3 buckets and objects, and retrieving and analyzing objects in your S3 buckets. If your account is the Macie administrator account for an organization, the policy also allows Macie to perform these actions on your behalf for member accounts in your organization.

If an S3 object is encrypted, the permissions policy for the Macie service-linked role typically grants Macie the permissions that it requires to decrypt the object. However, this depends on the type of encryption that was used. It can also depend on whether Macie is allowed to use the appropriate encryption key.

Topics

- Encryption options for Amazon S3 objects
- Allowing Macie to use a customer managed AWS KMS key

Encryption options for Amazon S3 objects

Amazon S3 supports multiple encryption options for S3 objects. For most of these options, Amazon Macie can decrypt an object by using the Macie service-linked role for your account. However, this depends on the type of encryption that was used to encrypt an object.

Server-side encryption with Amazon S3 managed keys (SSE-S3)

If an object is encrypted using server-side encryption with an Amazon S3 managed key (SSE-S3), Macie can decrypt the object.

To learn about this type of encryption, see Using server-side encryption with Amazon S3 managed keys in the *Amazon Simple Storage Service User Guide*.

Server-side encryption with AWS KMS keys (DSSE-KMS and SSE-KMS)

If an object is encrypted using dual-layer server-side encryption or server-side encryption with an AWS managed AWS KMS key (DSSE-KMS or SSE-KMS), Macie can decrypt the object.

If an object is encrypted using dual-layer server-side encryption or server-side encryption with a customer managed AWS KMS key (DSSE-KMS or SSE-KMS), Macie can decrypt the object only if you allow Macie to use the key. This is the case for objects that are encrypted with KMS keys managed entirely within AWS KMS and KMS keys in an external key store. If Macie isn't allowed to use the applicable KMS key, Macie can only store and report metadata for the object.

To learn about these types of encryption, see Using dual-layer server-side encryption with AWS KMS keys and Using server-side encryption with AWS KMS keys in the *Amazon Simple Storage Service User Guide*.

> **Tip**
>
> You can automatically generate a list of all the customer managed AWS KMS keys that Macie needs to access to analyze objects in S3 buckets for your account. To do this, run the AWS KMS Permission Analyzer script, which is available from the Amazon Macie Scripts repository on GitHub. The script can also generate an additional script of AWS Command Line Interface (AWS CLI) commands. You can optionally run those commands to update the requisite configuration settings and policies for KMS keys that you specify.

Server-side encryption with customer-provided keys (SSE-C)

If an object is encrypted using server-side encryption with a customer-provided key (SSE-C), Macie can't decrypt the object. Macie can only store and report metadata for the object.

To learn about this type of encryption, see Using server-side encryption with customer-provided keys in the *Amazon Simple Storage Service User Guide*.

Client-side encryption

If an object is encrypted using client-side encryption, Macie can't decrypt the object. Macie can only store and report metadata for the object. For example, Macie can report the size of the object and the tags that are associated with the object.

To learn about this type of encryption in the context of Amazon S3, see Protecting data by using client-side encryption in the *Amazon Simple Storage Service User Guide*.

You can filter your bucket inventory in Macie to determine which S3 buckets store objects that use certain types of encryption. You can also determine which buckets use certain types of server-side encryption by default when storing new objects. The following table provides examples of filters that you can apply to your bucket inventory to find this information.

To show buckets that...	Apply this filter...
Store objects that use SSE-C encryption	**Object count by encryption** is **Customer provided** and **From** = *1*
Store objects that use DSSE-KMS or SSE-KMS encryption	**Object count by encryption** is **AWS KMS managed** and **From** = *1*
Store objects that use SSE-S3 encryption	**Object count by encryption** is **Amazon S3 managed** and **From** = *1*
Store objects that use client-side encryption (or aren't encrypted)	**Object count by encryption** is **No encryption** and **From** = *1*
Encrypt new objects by default using DSSE-KMS encryption	**Default encryption** = **aws:kms:dsse**

To show buckets that...	Apply this filter...
Encrypt new objects by default using SSE-KMS encryption	**Default encryption** = **aws:kms**
Encrypt new objects by default using SSE-S3 encryption	**Default encryption** = **AES256**

If a bucket is configured to encrypt new objects by default using DSSE-KMS or SSE-KMS encryption, you can also determine which AWS KMS key is used. To do this, choose the bucket on the **S3 buckets** page. In the bucket details panel, under **Server-side encryption**, refer to the **AWS KMS key** field. This field shows the Amazon Resource Name (ARN) or unique identifier (key ID) for the key.

Allowing Macie to use a customer managed AWS KMS key

If an Amazon S3 object is encrypted using dual-layer server-side encryption or server-side encryption with a customer managed AWS KMS key (DSSE-KMS or SSE-KMS), Amazon Macie can decrypt the object only if it is allowed to use the key. How to provide this access depends on whether the account that owns the key also owns the S3 bucket that stores the object:

- If the same account owns the AWS KMS key and the bucket, a user of the account has to update the key's policy.
- If one account owns the AWS KMS key and a different account owns the bucket, a user of the account that owns the key has to allow cross-account access to the key.

This topic describes how to perform these tasks and provides examples for both scenarios. To learn more about allowing access to customer managed AWS KMS keys, see KMS key access and permissions in the *AWS Key Management Service Developer Guide*.

Allowing same-account access to a customer managed key

If the same account owns both the AWS KMS key and the S3 bucket, a user of the account has to add a statement to the policy for the key. The additional statement must allow the Macie service-linked role for the account to decrypt data by using the key. For detailed information about updating a key policy, see Changing a key policy in the *AWS Key Management Service Developer Guide*.

In the statement:

- The `Principal` element must specify the Amazon Resource Name (ARN) of the Macie service-linked role for the account that owns the AWS KMS key and the S3 bucket.

  If the account is in an opt-in AWS Region, the ARN must also include the appropriate Region code for the Region. For example, if the account is in the Middle East (Bahrain) Region, which has the Region code *me-south-1*, the `Principal` element must specify `arn:aws:iam::123456789012:role/aws-service-role/macie.me-south-1.amazonaws.com/AWSServiceRoleForAmazonMacie`, where *123456789012* is the account ID for the account. For a list of Region codes for the Regions where Macie is currently available, see Amazon Macie endpoints and quotas in the *AWS General Reference*.
- The `Action` array must specify the `kms:Decrypt` action. This is the only AWS KMS action that Macie must be allowed to perform to decrypt an S3 object that's encrypted with the key.

The following is an example of the statement to add to the policy for an AWS KMS key.

```
{
    "Sid": "Allow the Macie service-linked role to use the key",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::123456789012:role/aws-service-role/macie.amazonaws.com/
AWSServiceRoleForAmazonMacie"
    },
    "Action": [
        "kms:Decrypt"
    ],
    "Resource": "*"
}
```

In the preceding example:

- The `AWS` field in the `Principal` element specifies the ARN of the Macie service-linked role (`AWSServiceRoleForAmazonMacie`) for the account. It allows the Macie service-linked role to perform the action specified by the policy statement. *123456789012* is an example account ID. Replace this value with the account ID for the account that owns the KMS key and the S3 bucket.
- The `Action` array specifies the action that the Macie service-linked role is allowed to perform using the KMS key—decrypt ciphertext that's encrypted with the key.

Where you add this statement to a key policy depends on the structure and elements that the policy currently contains. When you add the statement, ensure that the syntax is valid. Key policies use JSON format. This means that you have to also add a comma before or after the statement, depending on where you add the statement to the policy.

Allowing cross-account access to a customer managed key

If one account owns the AWS KMS key (*key owner*) and a different account owns the S3 bucket (*bucket owner*), the key owner has to provide the bucket owner with cross-account access to the KMS key. To do this, the key owner first ensures that the key's policy allows the bucket owner to both use the key and create a grant for the key. The bucket owner then creates a grant for the key. A *grant* is a policy instrument that allows AWS principals to use KMS keys in cryptographic operations if the conditions specified by the grant are met. In this case, the grant delegates the relevant permissions to the Macie service-linked role for the bucket owner's account.

For detailed information about updating a key policy, see Changing a key policy in the *AWS Key Management Service Developer Guide*. To learn about grants, see Grants in AWS KMS in the *AWS Key Management Service Developer Guide*.

Step 1: Update the key policy

In the key policy, the key owner should ensure that the policy includes two statements:

- The first statement allows the bucket owner to use the key to decrypt data.
- The second statement allows the bucket owner to create a grant for the Macie service-linked role for their (the bucket owner's) account.

In the first statement, the `Principal` element must specify the ARN of the bucket owner's account. The `Action` array must specify the `kms:Decrypt` action. This is the only AWS KMS action that Macie must be allowed to perform to decrypt an object that's encrypted with the key. The following is an example of this statement in the policy for an AWS KMS key.

```
{
    "Sid": "Allow account 111122223333 to use the key",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::111122223333:root"
    },
    "Action": [
```

```
        "kms:Decrypt"
    ],
    "Resource": "*"
}
```

In the preceding example:

- The AWS field in the `Principal` element specifies the ARN of the bucket owner's account (*111122223333*). It allows the bucket owner to perform the action specified by the policy statement. *111122223333* is an example account ID. Replace this value with the account ID for the bucket owner's account.
- The `Action` array specifies the action that the bucket owner is allowed to perform using the KMS key—decrypt ciphertext that's encrypted with the key.

The second statement in the key policy allows the bucket owner to create a grant for the Macie service-linked role for their account. In this statement, the `Principal` element must specify the ARN of the bucket's owner's account. The `Action` array must specify the `kms:CreateGrant` action. A `Condition` element can filter access to the `kms:CreateGrant` action specified in the statement. The following is an example of this statement in the policy for an AWS KMS key.

```
{
    "Sid": "Allow account 111122223333 to create a grant",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::111122223333:root"
    },
    "Action": [
        "kms:CreateGrant"
    ],
    "Resource": "*",
    "Condition": {
        "StringEquals": {
            "kms:GranteePrincipal": "arn:aws:iam::111122223333:role/aws-service-role/
macie.amazonaws.com/AWSServiceRoleForAmazonMacie"
        }
    }
}
```

In the preceding example:

- The `AWS` field in the `Principal` element specifies the ARN of the bucket owner's account (*111122223333*). It allows the bucket owner to perform the action specified by the policy statement. *111122223333* is an example account ID. Replace this value with the account ID for the bucket owner's account.
- The `Action` array specifies the action that the bucket owner is allowed to perform on the KMS key—create a grant for the key.
- The `Condition` element uses the `StringEquals` condition operator and the `kms:GranteePrincipal` condition key to filter access to the action specified by the policy statement. In this case, the bucket owner can create a grant only for the specified `GranteePrincipal`, which is the ARN of the Macie service-linked role for their account. In that ARN, *111122223333* is an example account ID. Replace this value with the account ID for the bucket owner's account.

  If the bucket owner's account is in an opt-in AWS Region, also include the appropriate Region code in the ARN of the Macie service-linked role. For example, if the account is in the Middle East (Bahrain) Region, which has the Region code *me-south-1*, replace `macie.amazonaws.com` with `macie.me-south-1.amazonaws.com` in the ARN. For a list of Region codes for the Regions where Macie is currently available, see Amazon Macie endpoints and quotas in the *AWS General Reference*.

Where the key owner adds these statements to the key policy depends on the structure and elements that the policy currently contains. When the key owner adds the statements, they should ensure that the syntax is valid. Key policies use JSON format. This means that the key owner has to also add a comma before or after each statement, depending on where they add the statement to the policy.

Step 2: Create a grant

After the key owner updates the key policy as necessary, the bucket owner must create a grant for the key. The grant delegates the relevant permissions to the Macie service-linked role for their (the bucket owner's) account. Before the bucket owner creates the grant, they should verify that they're allowed to perform the `kms:CreateGrant` action for their account. This action allows them to add a grant to an existing, customer managed AWS KMS key.

To create the grant, the bucket owner can use the CreateGrant operation of the AWS Key Management Service API. When the bucket owner creates the grant, they should specify the following values for the required parameters:

- `KeyId` – The ARN of the KMS key. For cross-account access to a KMS key, this value must be an ARN. It can't be a key ID.
- `GranteePrincipal` – The ARN of the Macie service-linked role (`AWSServiceRoleForAmazonMacie`) for their account. This value should be `arn:aws:iam::`*`111122223333`*`:role/aws-service-role/macie.amazonaws.com/AWSServiceRoleForAmazonMacie`, where *`111122223333`* is the account ID for the bucket owner's account.

  If their account is in an opt-in Region, the ARN must include the appropriate Region code. For example, if their account is in the Middle East (Bahrain) Region, which has the Region code *me-south-1*, the ARN should be `arn:aws:iam::`*`111122223333`*`:role/aws-service-role/macie.me-south-1.amazonaws.com/AWSServiceRoleForAmazonMacie`, where *`111122223333`* is the account ID for the bucket owner's account.
- `Operations` – The AWS KMS decrypt action (`Decrypt`). This is the only AWS KMS action that Macie must be allowed to perform to decrypt an object that's encrypted with the KMS key.

To create a grant for a customer managed KMS key by using the AWS Command Line Interface (AWS CLI), run the create-grant command. The following example shows how. The example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws kms create-grant ^
--key-id arn:aws:kms:us-east-1:123456789012:key/1234abcd-12ab-34cd-56ef-1234567890ab ^
--grantee-principal arn:aws:iam::111122223333:role/aws-service-role/
macie.amazonaws.com/AWSServiceRoleForAmazonMacie ^
--operations "Decrypt"
```

Where:

- `key-id` specifies the ARN of the KMS key to apply the grant to.
- `grantee-principal` specifies the ARN of the Macie service-linked role for the account that's allowed to perform the action specified by the grant. This value should match the ARN specified by the `kms:GranteePrincipal` condition of the second statement in the key policy.
- `operations` specifies the action that the grant allows the specified principal to perform—decrypt ciphertext that's encrypted with the KMS key.

If the command runs successfully, you receive output similar to the following.

```
{
    "GrantToken": "<grant token>",
    "GrantId": "1a2b3c4d2f5e69f440bae30eaec9570bb1fb7358824f9ddfa1aa5a0dab1a59b2"
}
```

Where `GrantToken` is a unique, non-secret, variable-length, base64-encoded string that represents the grant that was created, and `GrantId` is the unique identifier for the grant.

Storing and retaining sensitive data discovery results

When you run a sensitive data discovery job or Amazon Macie performs automated sensitive data discovery, Macie creates an analysis record for each Amazon Simple Storage Service (Amazon S3) object that's included in the scope of the analysis. These records, referred to as a *sensitive data discovery results*, log details about the analysis that Macie performs on individual S3 objects. This includes objects that Macie doesn't detect sensitive data in, and therefore don't produce findings, and objects that Macie can't analyze due to errors or issues. If Macie detects sensitive data in an object, the record includes data from the corresponding finding as well as additional information. Sensitive data discovery results provide you with analysis records that can be helpful for data privacy and protection audits or investigations.

Macie stores your sensitive data discovery results for only 90 days. To access your results and enable long-term storage and retention of them, configure Macie to encrypt the results with an AWS Key Management Service (AWS KMS) key and store them in an S3 bucket. The bucket can serve as a definitive, long-term repository for all of your sensitive data discovery results. You can then optionally access and query the results in that repository.

This topic guides you through the process of using the AWS Management Console to configure a repository for your sensitive data discovery results. The configuration is a combination of an AWS KMS key that encrypts the results, an S3 general purpose bucket that stores the results, and Macie settings that specify which key and bucket to use. If you prefer to configure the Macie settings programmatically, you can use the PutClassificationExportConfiguration operation of the Amazon Macie API.

When you configure the settings in Macie, your choices apply only to the current AWS Region. If you're the Macie administrator for an organization, your choices apply only to your account. They don't apply to any associated member accounts. If you enable automated sensitive data discovery

or run sensitive data discovery jobs to analyze data for member accounts, Macie stores the sensitive data discovery results in the repository for your administrator account.

If you use Macie in multiple AWS Regions, configure the repository settings for each Region in which you use Macie. You can optionally store sensitive data discovery results for multiple Regions in the same S3 bucket. However, note the following requirements:

- To store the results for a Region that AWS enables by default for AWS accounts, such as the US East (N. Virginia) Region, you have to choose a bucket in a Region that's enabled by default. The results can't be stored in a bucket in an opt-in Region (Region that's disabled by default).
- To store the results for an opt-in Region, such as the Middle East (Bahrain) Region, you have to choose a bucket in that same Region or a Region that's enabled by default. The results can't be stored in a bucket in a different opt-in Region.

To determine whether a Region is enabled by default, see Enable or disable AWS Regions in your account in the *AWS Account Management User Guide*. In addition to the preceding requirements, also consider whether you want to retrieve samples of sensitive data that Macie reports in individual findings. To retrieve sensitive data samples from an affected S3 object, all of the following resources and data must be stored in the same Region: the affected object, the applicable finding, and the corresponding sensitive data discovery result.

Tasks

Before you begin: Learn key concepts

Amazon Macie automatically creates a sensitive data discovery result for each Amazon S3 object that it analyzes or attempts to analyze when you run a sensitive data discovery job or it performs automated sensitive data discovery. This includes:

- Objects that Macie detects sensitive data in, and therefore also produce sensitive data findings.
- Objects that Macie doesn't detect sensitive data in, and therefore don't produce sensitive data findings.

- Objects that Macie can't analyze due to errors or issues such as permissions settings or use of an unsupported file or storage format.

If Macie detects sensitive data in an S3 object, the sensitive data discovery result includes data from the corresponding sensitive data finding. It provides additional information too, such as the location of as many as 1,000 occurrences of each type of sensitive data that Macie found in the object. For example:

- The column and row number for a cell or field in a Microsoft Excel workbook, CSV file, or TSV file
- The path to a field or array in a JSON or JSON Lines file
- The line number for a line in a non-binary text file other than a CSV, JSON, JSON Lines, or TSV file—for example, an HTML, TXT, or XML file
- The page number for a page in an Adobe Portable Document Format (PDF) file
- The record index and the path to a field in a record in an Apache Avro object container or Apache Parquet file

If the affected S3 object is an archive file, such as a .tar or .zip file, the sensitive data discovery result also provides detailed location data for occurrences of sensitive data in individual files that Macie extracted from the archive. Macie doesn't include this information in sensitive data findings for archive files. To report location data, sensitive data discovery results use a standardized JSON schema.

A sensitive data discovery result doesn't include the sensitive data that Macie found. Instead, it provides you with an analysis record that can be helpful for audits or investigations.

Macie stores your sensitive data discovery results for 90 days. You can't access them directly on the Amazon Macie console or with the Amazon Macie API. Instead, follow the steps in this topic to configure Macie to encrypt your results with an AWS KMS key that you specify, and store the results in an S3 general purpose bucket that you also specify. Macie then writes the results to JSON Lines (.jsonl) files, adds the files to the bucket as GNU Zip (.gz) files, and encrypts the data using SSE-KMS encryption. As of November 8, 2023, Macie also signs the resulting S3 objects with a Hash-based Message Authentication Code (HMAC) AWS KMS key.

After you configure Macie to store your sensitive data discovery results in an S3 bucket, the bucket can serve as a definitive, long-term repository for the results. You can then optionally access and query the results in that repository.

Tips

For a detailed, instructional example of how you might query and use sensitive data discovery results to analyze and report potential data security risks, see the following blog post on the *AWS Security Blog*: How to query and visualize Macie sensitive data discovery results with Amazon Athena and Amazon QuickSight.

For samples of Amazon Athena queries that you can use to analyze sensitive data discovery results, visit the Amazon Macie Results Analytics repository on GitHub. This repository also provides instructions for configuring Athena to retrieve and decrypt your results, and scripts for creating tables for the results.

Step 1: Verify your permissions

Before you configure a repository for your sensitive data discovery results, verify that you have the permissions that you need to encrypt and store the results. To verify your permissions, use AWS Identity and Access Management (IAM) to review the IAM policies that are attached to your IAM identity. Then compare the information in those policies to the following list of actions that you must be allowed to perform to configure the repository.

Amazon Macie

For Macie, verify that you're allowed to perform the following action:

`macie2:PutClassificationExportConfiguration`

This action allows you to add or change the repository settings in Macie.

Amazon S3

For Amazon S3, verify that you're allowed to perform the following actions:

- `s3:CreateBucket`
- `s3:GetBucketLocation`
- `s3:ListAllMyBuckets`
- `s3:PutBucketAcl`
- `s3:PutBucketPolicy`
- `s3:PutBucketPublicAccessBlock`
- `s3:PutObject`

These actions allow you to access and configure an S3 general purpose bucket that can serve as the repository.

AWS KMS

To use the Amazon Macie console to add or change the repository settings, also verify that you're allowed to perform the following AWS KMS actions:

- `kms:DescribeKey`
- `kms:ListAliases`

These actions allow you to retrieve and display information about the AWS KMS keys for your account. You can then choose one of these keys to encrypt your sensitive data discovery results.

If you plan to create a new AWS KMS key to encrypt the data, you also need to be allowed to perform the following actions: `kms:CreateKey`, `kms:GetKeyPolicy`, and `kms:PutKeyPolicy`.

If you're not allowed to perform the requisite actions, ask your AWS administrator for assistance before you proceed to the next step.

Step 2: Configure an AWS KMS key

After you verify your permissions, determine which AWS KMS key you want Macie to use to encrypt your sensitive data discovery results. The key must be a customer managed, symmetric encryption KMS key that's enabled in the same AWS Region as the S3 bucket where you want to store the results.

The key can be an existing AWS KMS key from your own account, or an existing AWS KMS key that another account owns. If you want to use a new KMS key, create the key before proceeding. If you want to use an existing key that another account owns, obtain the Amazon Resource Name (ARN) of the key. You'll need to enter this ARN when you configure the repository settings in Macie. For information about creating and reviewing the settings for KMS keys, see the AWS Key Management Service Developer Guide.

> **Note**
>
> The key can be an AWS KMS key in an external key store. However, the key might then be slower and less reliable than a key that's managed entirely within AWS KMS. You can reduce this risk by storing your sensitive data discovery results in an S3 bucket that's

configured to use the key as an S3 Bucket Key. Doing so reduces the number of AWS KMS requests that must be made to encrypt your sensitive data discovery results.
For information about using KMS keys in external key stores, see External key stores in the *AWS Key Management Service Developer Guide*. For information about using S3 Bucket Keys, see Reducing the cost of SSE-KMS with Amazon S3 Bucket Keys in the *Amazon Simple Storage Service User Guide*.

After you determine which KMS key you want Macie to use, give Macie permission to use the key. Otherwise, Macie won't be able to encrypt or store your results in the repository. To give Macie permission to use the key, update the key policy for the key. For detailed information about key policies and managing access to KMS keys, see Key policies in AWS KMS in the *AWS Key Management Service Developer Guide*.

To update the key policy

1. Open the AWS KMS console at https://console.aws.amazon.com/kms.
2. To change the AWS Region, use the Region selector in the upper-right corner of the page.
3. Choose the key that you want Macie to use to encrypt your sensitive data discovery results.
4. On the **Key policy** tab, choose **Edit**.
5. Copy the following statement to your clipboard, and then add it to the policy:

```
{
    "Sid": "Allow Macie to use the key",
    "Effect": "Allow",
    "Principal": {
        "Service": "macie.amazonaws.com"
    },
    "Action": [
        "kms:GenerateDataKey",
        "kms:Encrypt"
    ],
    "Resource": "*",
    "Condition": {
        "StringEquals": {
            "aws:SourceAccount": "111122223333"
         },
         "ArnLike": {
             "aws:SourceArn": [
```

```
                "arn:aws:macie2:us-east-1:111122223333:export-configuration:*",
                "arn:aws:macie2:us-east-1:111122223333:classification-job/*"
            ]
        }
    }
}
```

> **Note**
>
> When you add the statement to the policy, make sure that the syntax is valid. Policies use JSON format. This means that you need to also add a comma before or after the statement, depending on where you add the statement to the policy. If you add the statement as the last statement, add a comma after the closing curly brace for the preceding statement. If you add it as the first statement or between two existing statements, add a comma after the closing curly brace for the statement.

6. Update the statement with the correct values for your environment:

- In the `Condition` fields, replace the placeholder values, where:
  - *111122223333* is the account ID for your AWS account.
  - *us-east-1* is the Region code for the AWS Region in which you're using Macie and want to allow Macie to use the key.

    If you use Macie in multiple Regions and want to allow Macie to use the key in additional Regions, add `aws:SourceArn` conditions for each additional Region. For example:

    ```
    "aws:SourceArn": [
        "arn:aws:macie2:us-east-1:111122223333:export-configuration:*",
        "arn:aws:macie2:us-east-1:111122223333:classification-job/*",
        "arn:aws:macie2:us-west-2:111122223333:export-configuration:*",
        "arn:aws:macie2:us-west-2:111122223333:classification-job/*"
    ]
    ```

    Alternatively, you can allow Macie to use the key in all Regions. To do this, replace the placeholder value with the wildcard character (*). For example:

    ```
    "aws:SourceArn": [
        "arn:aws:macie2:*:111122223333:export-configuration:*",
        "arn:aws:macie2:*:111122223333:classification-job/*"
    ```

```
]
```

- If you're using Macie in an opt-in Region, add the appropriate Region code to the value for the `Service` field. For example, if you're using Macie in the Middle East (Bahrain) Region, which has the Region code *me-south-1*, replace `macie.amazonaws.com` with `macie.me-south-1.amazonaws.com`.

  For a list of Regions where Macie is currently available and the Region code for each one, see Amazon Macie endpoints and quotas in the *AWS General Reference*.

Note that the `Condition` fields use two IAM global condition keys:

- aws:SourceAccount – This condition allows Macie to perform the specified actions only for your account. More specifically, it determines which account can perform the specified actions for the resources and actions specified by the `aws:SourceArn` condition.

  To allow Macie to perform the specified actions for additional accounts, add the account ID for each additional account to this condition. For example:

  ```
  "aws:SourceAccount": [111122223333,444455556666]
  ```

- aws:SourceArn – This condition prevents other AWS services from performing the specified actions. It also prevents Macie from using the key while performing other actions for your account. In other words, it allows Macie to encrypt S3 objects with the key only if: the objects are sensitive data discovery results, and the results are for automated sensitive data discovery or sensitive data discovery jobs created by the specified account in the specified Region.

  To allow Macie to perform the specified actions for additional accounts, add ARNs for each additional account to this condition. For example:

  ```
  "aws:SourceArn": [
      "arn:aws:macie2:us-east-1:111122223333:export-configuration:*",
      "arn:aws:macie2:us-east-1:111122223333:classification-job/*",
      "arn:aws:macie2:us-east-1:444455556666:export-configuration:*",
      "arn:aws:macie2:us-east-1:444455556666:classification-job/*"
  ]
  ```

The accounts specified by the `aws:SourceAccount` and `aws:SourceArn` conditions should match.

These conditions help prevent Macie from being used as a [confused deputy] during transactions with AWS KMS. Although we don't recommend it, you can remove these conditions from the statement.

7. When you finish adding and updating the statement, choose **Save changes**.

Step 3: Choose an S3 bucket

After you verify your permissions and configure the AWS KMS key, you're ready to specify which S3 bucket you want to use as the repository for your sensitive data discovery results. You have two options:

- **Use a new S3 bucket that Macie creates** – If you choose this option, Macie automatically creates a new S3 general purpose bucket in the current AWS Region for your discovery results. Macie also applies a bucket policy to the bucket. The policy allows Macie to add objects to the bucket. It also requires the objects to be encrypted with the AWS KMS key that you specify, using SSE-KMS encryption. To review the policy, choose **View policy** on the Amazon Macie console after you specify a name for the bucket and the KMS key to use.
- **Use an existing S3 bucket that you create** – If you prefer to store your discovery results in a particular S3 bucket that you create, create the bucket before you proceed. The bucket must be a general purpose bucket. In addition, the bucket's settings and policy must allow Macie to add objects to the bucket. This topic explains which settings to check and how to update the policy. It also provides examples of the statements to add to the policy.

The following sections provide instructions for each option. Choose the section for the option that you want.

Use a new S3 bucket that Macie creates

If you prefer to use a new S3 bucket that Macie creates for you, the final step in the process is to configure the repository settings in Macie.

To configure the repository settings in Macie

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, under **Settings**, choose **Discovery results**.
3. Under **Repository for sensitive data discovery results**, choose **Create bucket**.
4. In the **Create a bucket** box, enter a name for the bucket.

   The name must be unique across all S3 buckets. In addition, the name can consist only of lowercase letters, numbers, dots (.), and hyphens (-). For additional naming requirements, see Bucket naming rules in the *Amazon Simple Storage Service User Guide*.
5. Expand the **Advanced** section.
6. (Optional) To specify a prefix to use in the path to a location in the bucket, enter the prefix in the **Data discovery result prefix** box.

   When you enter a value, Macie updates the example below the box to show the path to the bucket location where it will store your discovery results.
7. For **Block all public access**, choose **Yes** to enable all block public access settings for the bucket.

   For information about these settings, see Blocking public access to your Amazon S3 storage in the *Amazon Simple Storage Service User Guide*.
8. Under **Encryption settings**, specify the AWS KMS key that you want Macie to use to encrypt the results:

   - To use a key from your own account, choose **Select a key from your account**. Then, in the **AWS KMS key** list, choose the key to use. The list displays customer managed, symmetric encryption KMS keys for your account.
   - To use a key that another account owns, choose **Enter the ARN of a key from another account**. Then, in the **AWS KMS key ARN** box, enter the Amazon Resource Name (ARN) of the key to use—for example, **`arn:aws:kms:us-east-1:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab`**.
9. When you finish entering the settings, choose **Save**.

   Macie tests the settings to verify that they're correct. If any settings are incorrect, Macie displays an error message to help you address the issue.

After you save the repository settings, Macie adds existing discovery results for the preceding 90 days to the repository. Macie also starts adding new discovery results to the repository.

Use an existing S3 bucket that you create

If you prefer to store your sensitive data discovery results in a particular S3 bucket that you create, create and configure the bucket before you configure the settings in Macie. When you create the bucket, note the following requirements:

- The bucket must be a general purpose bucket. It can't be another type of bucket, such as a directory bucket.
- To store your discovery results for a Region that's enabled by default for AWS accounts, such as the US East (N. Virginia) Region, the bucket has to be in a Region that's enabled by default. The results can't be stored in a bucket in an opt-in Region (Region that's disabled by default).
- To store your discovery results for an opt-in Region, such as the Middle East (Bahrain) Region, the bucket has to be in the same Region or a Region that's enabled by default. The results can't be stored in a bucket in a different opt-in Region.

To determine whether a Region is enabled by default, see Enable or disable AWS Regions in your account in the *AWS Account Management User Guide*.

After you create the bucket, update the bucket's policy to allow Macie to retrieve information about the bucket and add objects to the bucket. You can then configure the settings in Macie.

To update the bucket policy for the bucket

1. Open the Amazon S3 console at https://console.aws.amazon.com/s3/.
2. Choose the bucket that you want to store your discovery results in.
3. Choose the **Permissions** tab.
4. In the **Bucket policy** section, choose **Edit**.
5. Copy the following example policy to your clipboard:

   JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
```

```
            "Sid": "Allow Macie to use the GetBucketLocation operation",
            "Effect": "Allow",
            "Principal": {
                "Service": "macie.amazonaws.com"
            },
            "Action": "s3:GetBucketLocation",
            "Resource": "arn:aws:s3:::amzn-s3-demo-bucket",
            "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "111122223333"
                },
                "ArnLike": {
                    "aws:SourceArn": [
                        "arn:aws:macie2:us-east-1:111122223333:export-
configuration:*",
                        "arn:aws:macie2:us-
east-1:111122223333:classification-job/*"
                    ]
                }
            }
        },
        {
            "Sid": "Allow Macie to add objects to the bucket",
            "Effect": "Allow",
            "Principal": {
                "Service": "macie.amazonaws.com"
            },
            "Action": "s3:PutObject",
            "Resource": "arn:aws:s3:::amzn-s3-demo-bucket/[optional
 prefix/]*",
            "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "111122223333"
                },
                "ArnLike": {
                    "aws:SourceArn": [
                        "arn:aws:macie2:us-east-1:111122223333:export-
configuration:*",
                        "arn:aws:macie2:us-
east-1:111122223333:classification-job/*"
                    ]
                }
            }
        },
```

```
        {
            "Sid": "Deny unencrypted object uploads. This is optional",
            "Effect": "Deny",
            "Principal": {
                "Service": "macie.amazonaws.com"
            },
            "Action": "s3:PutObject",
            "Resource": "arn:aws:s3:::amzn-s3-demo-bucket/[optional
prefix/]*",
            "Condition": {
                "StringNotEquals": {
                    "s3:x-amz-server-side-encryption": "aws:kms"
                }
            }
        },
        {
            "Sid": "Deny incorrect encryption headers. This is optional",
            "Effect": "Deny",
            "Principal": {
                "Service": "macie.amazonaws.com"
            },
            "Action": "s3:PutObject",
            "Resource": "arn:aws:s3:::amzn-s3-demo-bucket/[optional
prefix/]*",
            "Condition": {
                "StringNotEquals": {
                    "s3:x-amz-server-side-encryption-aws-kms-key-id":
"arn:aws:kms:us-east-1:111122223333:key/KMSKeyId"
                }
            }
        },
        {
            "Sid": "Deny non-HTTPS access",
            "Effect": "Deny",
            "Principal": "*",
            "Action": "s3:*",
            "Resource": "arn:aws:s3:::amzn-s3-demo-bucket/*",
            "Condition": {
                "Bool": {
                    "aws:SecureTransport": "false"
                }
            }
        }
    ]
```

```
}
```

6. Paste the example policy in the **Bucket policy** editor on the Amazon S3 console.
7. Update the example policy with the correct values for your environment:

   - In the optional statement that denies incorrect encryption headers:
     - Replace *amzn-s3-demo-bucket* with the name of the bucket. To also specify a prefix for a path to a location in the bucket, replace *[optional prefix/]* with the prefix. Otherwise, remove the *[optional prefix/]* placeholder value.
     - In the `StringNotEquals` condition, replace *arn:aws:kms:us-east-1:111122223333:key/KMSKeyId* with the Amazon Resource Name (ARN) of the AWS KMS key to use for encryption of your discovery results.
   - In all other statements, replace the placeholder values, where:
     - *amzn-s3-demo-bucket* is the name of the bucket.
     - *[optional prefix/]* is the prefix for a path to a location in the bucket. Remove this placeholder value if you don't want to specify a prefix.
     - *111122223333* is the account ID for your AWS account.
     - *us-east-1* is the Region code for the AWS Region in which you're using Macie and want to allow Macie to add discovery results to the bucket.

       If you use Macie in multiple Regions and want to allow Macie to add results to the bucket for additional Regions, add `aws:SourceArn` conditions for each additional Region. For example:

       ```
       "aws:SourceArn": [
           "arn:aws:macie2:us-east-1:111122223333:export-configuration:*",
           "arn:aws:macie2:us-east-1:111122223333:classification-job/*",
           "arn:aws:macie2:us-west-2:111122223333:export-configuration:*",
           "arn:aws:macie2:us-west-2:111122223333:classification-job/*"
       ]
       ```

       Alternatively, you can allow Macie to add results to the bucket for all Regions in which you use Macie. To do this, replace the placeholder value with the wildcard character (*). For example:

       ```
       "aws:SourceArn": [
           "arn:aws:macie2:*:111122223333:export-configuration:*",
       ```

```
    "arn:aws:macie2:*:111122223333:classification-job/*"
]
```

- If you're using Macie in an opt-in Region, add the appropriate Region code to the value for the `Service` field in each statement that specifies the Macie service principal. For example, if you're using Macie in the Middle East (Bahrain) Region, which has the Region code *me-south-1*, replace `macie.amazonaws.com` with `macie.me-south-1.amazonaws.com` in each applicable statement.

  For a list of Regions where Macie is currently available and the Region code for each one, see Amazon Macie endpoints and quotas in the *AWS General Reference*.

Note that the example policy includes statements that allow Macie to determine which Region the bucket resides in (`GetBucketLocation`) and add objects to the bucket (`PutObject`). These statements define conditions that use two IAM global condition keys:

- aws:SourceAccount – This condition allows Macie to add sensitive data discovery results to the bucket only for your account. It prevents Macie from adding discovery results for other accounts to the bucket. More specifically, the condition specifies which account can use the bucket for the resources and actions specified by the `aws:SourceArn` condition.

  To store results for additional accounts in the bucket, add the account ID for each additional account to this condition. For example:

  ```
  "aws:SourceAccount": [111122223333,444455556666]
  ```

- aws:SourceArn – This condition restricts access to the bucket based on the source of the objects that are being added to the bucket. It prevents other AWS services from adding objects to the bucket. It also prevents Macie from adding objects to the bucket while performing other actions for your account. More specifically, the condition allows Macie to add objects to the bucket only if: the objects are sensitive data discovery results, and the results are for automated sensitive data discovery or sensitive data discovery jobs created by the specified account in the specified Region.

  To allow Macie to perform the specified actions for additional accounts, add ARNs for each additional account to this condition. For example:

  ```
  "aws:SourceArn": [
      "arn:aws:macie2:us-east-1:111122223333:export-configuration:*",
  ```

```
    "arn:aws:macie2:us-east-1:111122223333:classification-job/*",
    "arn:aws:macie2:us-east-1:444455556666:export-configuration:*",
    "arn:aws:macie2:us-east-1:444455556666:classification-job/*"
]
```

The accounts specified by the `aws:SourceAccount` and `aws:SourceArn` conditions should match.

Both conditions help prevent Macie from being used as a confused deputy during transactions with Amazon S3. Although we don't recommend it, you can remove these conditions from the bucket policy.

8. When you finish updating the bucket policy, choose **Save changes**.

You can now configure the repository settings in Macie.

To configure the repository settings in Macie

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, under **Settings**, choose **Discovery results**.
3. Under **Repository for sensitive data discovery results**, choose **Existing bucket**.
4. For **Choose a bucket**, select the bucket that you want to store your discovery results in.
5. To specify a prefix for a path to a location in the bucket, expand the **Advanced** section. Then, for **Data discovery result prefix**, enter the prefix.

   When you enter a value, Macie updates the example below the box to show the path to the bucket location where it will store your discovery results.
6. Under **Encryption settings**, specify the AWS KMS key that you want Macie to use to encrypt the results:

   - To use a key from your own account, choose **Select a key from your account**. Then, in the **AWS KMS key** list, choose the key to use. The list displays customer managed, symmetric encryption KMS keys for your account.
   - To use a key that another account owns, choose **Enter the ARN of a key from another account**. Then, in the **AWS KMS key ARN** box, enter the ARN of the key to use—for example, **`arn:aws:kms:us-east-1:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab`**.

7. When you finish entering the settings, choose **Save**.

   Macie tests the settings to verify that they're correct. If any settings are incorrect, Macie displays an error message to help you address the issue.

After you save the repository settings, Macie adds existing discovery results for the preceding 90 days to the repository. Macie also starts adding new discovery results to the repository.

> **Note**
>
> If you subsequently change the **Data discovery result prefix** setting, also update the bucket policy in Amazon S3. Policy statements that specify the previous prefix must specify the new prefix. Otherwise, Macie won't be allowed to add your discovery results to the bucket.

> **Tip**
>
> To reduce server-side encryption costs, also configure the S3 bucket to use an S3 Bucket Key, and specify the AWS KMS key that you configured for encryption of your sensitive data discovery results. Use of an S3 Bucket Key reduces the number of calls to AWS KMS, which can reduce AWS KMS request costs. If the KMS key is in an external key store, use of an S3 Bucket Key can also minimize the performance impact of using the key. To learn more, see Reducing the cost of SSE-KMS with Amazon S3 Bucket Keys in the *Amazon Simple Storage Service User Guide*.

Supported storage classes and formats

To help you discover sensitive data in your Amazon Simple Storage Service (Amazon S3) data estate, Amazon Macie supports most Amazon S3 storage classes and a wide variety of file and storage formats. This support applies to the use of managed data identifiers and the use of custom data identifiers to analyze S3 objects.

For Macie to analyze an S3 object, the object has to be stored in an Amazon S3 general purpose bucket using a supported storage class. The object also has to use a supported file or storage format. The topics in this section list the storage classes and the file and storage formats that Macie currently supports.

Tip

Although Macie is optimized for Amazon S3, you can use it to discover sensitive data in resources that you currently store elsewhere. You can do this by moving the data to Amazon S3 temporarily or permanently. For example, export Amazon Relational Database Service or Amazon Aurora snapshots to Amazon S3 in Apache Parquet format. Or export an Amazon DynamoDB table to Amazon S3. You can then create a sensitive data discovery job to analyze the data in Amazon S3.

Topics

Supported Amazon S3 storage classes

For sensitive data discovery, Amazon Macie supports the following Amazon S3 storage classes:

- Reduced Redundancy (RRS)
- S3 Glacier Instant Retrieval
- S3 Intelligent-Tiering
- S3 One Zone-Infrequent Access (S3 One Zone-IA)
- S3 Standard
- S3 Standard-Infrequent Access (S3 Standard-IA)

Macie doesn't analyze S3 objects that use other Amazon S3 storage classes, such as S3 Glacier Deep Archive or S3 Express One Zone. In addition, Macie doesn't analyze objects that are stored in S3 directory buckets.

If you configure a sensitive data discovery job to analyze S3 objects that don't use a supported Amazon S3 storage class, Macie skips those objects when the job runs. Macie doesn't attempt to retrieve or analyze data in the objects—the objects are treated as *unclassifiable objects*. An *unclassifiable object* is an object that doesn't use a supported storage class or a supported file or storage format. Macie analyzes only those objects that use a supported storage class and a supported file or storage format.

Similarly, if you configure Macie to perform automated sensitive data discovery, unclassifiable objects aren't eligible for selection and analysis. Macie selects only those objects that use a supported Amazon S3 storage class and a supported file or storage format.

To identify S3 buckets that store unclassifiable objects, you can filter your S3 bucket inventory. For each bucket in your inventory, there are fields that report the number and total storage size of unclassifiable objects in the bucket.

For detailed information about the storage classes that Amazon S3 provides, see Using Amazon S3 storage classes in the *Amazon Simple Storage Service User Guide*.

Supported file and storage formats

When Amazon Macie analyzes an S3 object, Macie retrieves the latest version of the object from Amazon S3, and then performs a deep inspection of the object's contents. This inspection factors the file or storage format of the data. Macie can analyze data in many different formats, including commonly used compression and archive formats.

When Macie analyzes data in a compressed or archive file, Macie inspects both the full file and the contents of the file. To inspect the file's contents, Macie decompresses the file, and then inspects each extracted file that uses a supported format. Macie can do this for as many as 1,000,000 files and up to a nested depth of 10 levels. For information about additional quotas that apply to sensitive data discovery, see Quotas for Macie.

The following table lists and describes the types of file and storage formats that Macie can analyze to detect sensitive data. For each supported type, the table also lists the applicable file name extensions.

File or storage type	Description	File name extensions
Big data	Apache Avro object containers and Apache Parquet files	.avro, .parquet
Compression or archive	GNU Zip compressed archives, TAR archives, and ZIP compressed archives	.gz, .gzip, .tar, .zip
Document	Adobe Portable Document Format files, Microsoft Excel	.doc, .docx, .pdf, .xls, .xlsx

File or storage type	Description	File name extensions
	workbooks, and Microsoft Word documents	
Email message	Electronic mail files whose contents comply with the requirements specified by an IETF RFC for electronic mail messages, such as RFC 2822	.eml
Text	Non-binary text files. Examples are: comma-separated values (CSV) files, Extensible Markup Language (XML) files, Hypertext Markup Language (HTML) files, JavaScript Object Notation (JSON) files, JSON Lines files, plaintext documents, tab-separated values (TSV) files, and YAML files	Depending on the type of non-binary text file: .csv, .htm, .html, .json, .jsonl, .tsv, . and others

Macie doesn't analyze data in images, or audio, video, and other types of multimedia content.

If you configure a sensitive data discovery job to analyze S3 objects that don't use a supported file or storage format, Macie skips those objects when the job runs. Macie doesn't attempt to retrieve or analyze data in the objects—the objects are treated as *unclassifiable objects*. An *unclassifiable object* is an object that doesn't use a supported Amazon S3 storage class or a supported file or storage format. Macie analyzes only those objects that use a supported storage class and a supported file or storage format.

Similarly, if you configure Macie to perform automated sensitive data discovery, unclassifiable objects aren't eligible for selection and analysis. Macie selects only those objects that use a supported Amazon S3 storage class and a supported file or storage format.

To identify S3 buckets that store unclassifiable objects, you can filter your S3 bucket inventory. For each bucket in your inventory, there are fields that report the number and total storage size of unclassifiable objects in the bucket.

Reviewing and analyzing Macie findings

Amazon Macie generates findings when it detects potential policy violations or issues with the security or privacy of your Amazon Simple Storage Service (Amazon S3) general purpose buckets or it detects sensitive data in S3 objects. A *finding* is a detailed report of a potential issue or sensitive data that Macie found. Each finding provides a severity rating, information about the affected resource, and additional details, such as when and how Macie found the issue or data. Macie stores your policy and sensitive data findings for 90 days.

You can review, analyze, and manage findings in the following ways.

Amazon Macie console

The **Findings** pages on the Amazon Macie console list your findings and provide detailed information for individual findings. These pages also provide options for grouping, filtering, and sorting findings, and for creating and managing suppression rules. Suppression rules can help you streamline your analysis of findings.

Amazon Macie API

With the Amazon Macie API, you can query and retrieve findings data by using an AWS command line tool or an AWS SDK, or by sending HTTPS requests directly to Macie. To query the data, you submit a request to the Amazon Macie API and use supported parameters to specify which findings you want to retrieve. After you submit your request, Macie returns the results in a JSON response. You can then pass the results to another service or application for deeper analysis, long-term storage, or reporting. For more information, see the Amazon Macie API Reference.

Amazon EventBridge

To further support integration with other services and systems, such as monitoring or event management systems, Macie publishes findings to Amazon EventBridge as events. EventBridge, formerly Amazon CloudWatch Events, is a serverless event bus service that can deliver a stream of real-time data from your own applications, software as a service (SaaS) applications, and AWS services such as Macie. It can route that data to targets such as AWS Lambda functions, Amazon Simple Notification Service topics, and Amazon Kinesis streams for additional, automated processing. Use of EventBridge also helps ensure longer-term retention of findings data. To learn more about EventBridge, see the Amazon EventBridge User Guide.

Macie automatically publishes events to EventBridge for new findings. It also publishes events automatically for subsequent occurrences of existing policy findings. Because the findings data is structured as EventBridge events, you can more easily monitor, analyze, and act upon findings by using other services and tools. For example, you might use EventBridge to automatically send specific types of new findings to an AWS Lambda function that, in turn, processes and sends the data to your security incident and event management (SIEM) system. If you integrate AWS User Notifications with Macie, you can also use the events to be notified of findings automatically through delivery channels that you specify. To learn about using EventBridge events to monitor and process findings, see Processing findings with Amazon EventBridge.

AWS Security Hub

For additional, broader analysis of your organization's security posture, you can also publish findings to AWS Security Hub. Security Hub is a service that collects security data from AWS services and supported AWS Partner Network security solutions to provide you with a comprehensive view of your security state across your AWS environment. Security Hub also helps you check your environment against security industry standards and best practices. To learn more about Security Hub, see the AWS Security Hub User Guide. To learn about using Security Hub to evaluate and process findings, see Evaluating findings with AWS Security Hub.

In addition to findings, Macie creates sensitive data discovery results for S3 objects that it analyzes to discover sensitive data. A *sensitive data discovery result* is a record that logs details about the analysis of an object. This includes objects that Macie doesn't find sensitive data in, and therefore don't produce findings, and objects that Macie can't analyze due to errors or issues. Sensitive data discovery results provide you with analysis records that can be helpful for data privacy and protection audits or investigations. You can't access sensitive data discovery results directly on the Amazon Macie console or with the Amazon Macie API. Instead, you configure Macie to store the results in an S3 bucket. You can then optionally access and query the results in that bucket. To learn how to configure Macie to store the results, see Storing and retaining sensitive data discovery results.

Topics

- Investigating sensitive data with Macie findings
- Suppressing Macie findings

Types of Macie findings

Amazon Macie generates two categories of findings: *policy findings* and *sensitive data findings*. A *policy finding* is a detailed report of a potential policy violation or issue with the security or privacy of an Amazon Simple Storage Service (Amazon S3) general purpose bucket. Macie generates policy findings as part of its ongoing activities to evaluate and monitor your general purpose buckets for security and access control. A *sensitive data finding* is a detailed report of sensitive data that Macie detected in an S3 object. Macie generates sensitive data findings as part of the activities that it performs when you run sensitive data discovery jobs or it performs automated sensitive data discovery.

Within each category, there are specific types. A finding's type provides insight into the nature of the issue or sensitive data that Macie found. A finding's details provide a severity rating, information about the affected resource, and additional information, such as when and how Macie found the issue or sensitive data. The severity and details of each finding vary depending on the type and nature of the finding.

Topics

- Types of policy findings
- Types of sensitive data findings

> **Tip**
>
> To explore and learn about the different categories and types of findings that Macie can generate, create sample findings. Sample findings use example data and placeholder values to demonstrate the kinds of information that each type of finding might contain.

Types of policy findings

Amazon Macie generates a policy finding when the policies or settings for an S3 general purpose bucket are changed in a way that reduces the security or privacy of the bucket and the bucket's

objects. For information about how Macie detects and evaluates these changes, see How Macie monitors Amazon S3 data security.

Note that Macie generates a policy finding only if the change occurs after you enable Macie for your AWS account. For example, if block public access settings are disabled for an S3 bucket after you enable Macie, Macie generates a **Policy:IAMUser/S3BlockPublicAccessDisabled** finding for the bucket. If block public access settings were disabled for a bucket when you enabled Macie and they continue to be disabled, Macie doesn't generate a **Policy:IAMUser/S3BlockPublicAccessDisabled** finding for the bucket.

If Macie detects a subsequent occurrence of an existing policy finding, Macie updates the existing finding by adding details about the subsequent occurrence and incrementing the count of occurrences. Macie stores policy findings for 90 days.

Macie can generate the following types of policy findings for an S3 general purpose bucket.

Policy:IAMUser/S3BlockPublicAccessDisabled

All bucket-level block public access settings were disabled for the bucket. Public access to the bucket is controlled by the block public access settings for the account, access control lists (ACLs), the bucket policy for the bucket, and other settings and policies that apply to the bucket.

To investigate the finding, start by reviewing the bucket's details in Macie. The details include a breakdown of the bucket's public access settings. For detailed information about the settings, see Access control and Blocking public access to your Amazon S3 storage in the *Amazon Simple Storage Service User Guide*.

Policy:IAMUser/S3BucketEncryptionDisabled

Default encryption settings for the bucket were reset to default Amazon S3 encryption behavior, which is to encrypt new objects automatically with an Amazon S3 managed key.

Starting January 5, 2023, Amazon S3 automatically applies server-side encryption with Amazon S3 managed keys (SSE-S3) as the base level of encryption for objects that are added to buckets. You can optionally configure a bucket's default encryption settings to instead use server-side encryption with an AWS KMS key (SSE-KMS) or dual-layer server-side encryption with an AWS KMS key (DSSE-KMS). If Macie generated this type of finding prior to January 5, 2023, the finding indicates that default encryption settings were disabled for the affected bucket. This meant that the bucket's settings didn't specify default server-side encryption behavior for new

objects. The ability to disable default encryption settings for a bucket is no longer supported by Amazon S3.

To learn about default encryption settings and options for S3 buckets, see Setting default server-side encryption behavior for S3 buckets in the *Amazon Simple Storage Service User Guide*.

Policy:IAMUser/S3BucketPublic

An ACL or bucket policy for the bucket was changed to allow access by anonymous users or all authenticated AWS Identity and Access Management (IAM) identities.

To investigate the finding, start by reviewing the bucket's details in Macie. The details include a breakdown of the bucket's public access settings. For detailed information about ACLs, bucket policies, and access settings for S3 buckets, see Access control in the *Amazon Simple Storage Service User Guide*.

Policy:IAMUser/S3BucketReplicatedExternally

Replication was enabled and configured to replicate objects from the bucket to a bucket for an AWS account that's external to (not part of) your organization. An *organization* is a set of Macie accounts that are centrally managed as a group of related accounts through AWS Organizations or by Macie invitation.

Under certain conditions, Macie might generate this type of finding for a bucket that isn't configured to replicate objects to a bucket for an external AWS account. This can occur if the destination bucket was created in a different AWS Region during the preceding 24 hours, after Macie retrieved bucket and object metadata from Amazon S3 as part of the daily refresh cycle.

To investigate the finding, start by refreshing your inventory data in Macie. Then review the bucket's details. The details indicate whether the bucket is configured to replicate objects to other buckets. If the bucket is configured to do this, the details include the account ID for each account that owns a destination bucket. For detailed information about replication settings for S3 buckets, see Replicating objects in the *Amazon Simple Storage Service User Guide*.

Policy:IAMUser/S3BucketSharedExternally

An ACL or bucket policy for the bucket was changed to allow the bucket to be shared with an AWS account that's external to (not part of) your organization. An *organization* is a set of Macie accounts that are centrally managed as a group of related accounts through AWS Organizations or by Macie invitation.

In certain cases, Macie might generate this type of finding for a bucket that isn't shared with an external AWS account. This can occur if Macie isn't able to fully evaluate the relationship

between the `Principal` element in the bucket's policy and certain AWS global condition context keys or Amazon S3 condition keys in the `Condition` element of the policy. This can be the case for the following condition keys: `aws:PrincipalAccount`, `aws:PrincipalArn`, `aws:PrincipalOrgID`, `aws:PrincipalOrgPaths`, `aws:PrincipalTag`, `aws:PrincipalType`, `aws:SourceAccount`, `aws:SourceArn`, `aws:SourceIp`, `aws:SourceOrgID`, `aws:SourceOrgPaths`, `aws:SourceVpc`, `aws:SourceVpce`, `aws:userid`, `s3:DataAccessPointAccount`, and `s3:DataAccessPointArn`. We recommend that you review the bucket's policy to determine whether this access is intended and safe.

To learn about ACLs and bucket policies for S3 buckets, see Access control in the *Amazon Simple Storage Service User Guide*.

Policy:IAMUser/S3BucketSharedWithCloudFront

The bucket policy for the bucket was changed to allow the bucket to be shared with an Amazon CloudFront origin access identity (OAI), a CloudFront origin access control (OAC), or both a CloudFront OAI and a CloudFront OAC. A CloudFront OAI or OAC allows users to access a bucket's objects through one or more specified CloudFront distributions.

To learn about CloudFront OAIs and OACs, see Restricting access to an Amazon S3 origin in the *Amazon CloudFront Developer Guide*.

Note

In certain cases, Macie generates a **Policy:IAMUser/S3BucketSharedExternally** finding instead of a **Policy:IAMUser/S3BucketSharedWithCloudFront** finding for a bucket. These cases are:

- The bucket is shared with an AWS account that's external to your organization, in addition to a CloudFront OAI or OAC.
- The bucket's policy specifies a canonical user ID, instead of the Amazon Resource Name (ARN), of a CloudFront OAI.

This produces a higher severity policy finding for the bucket.

Types of sensitive data findings

Amazon Macie generates a sensitive data finding when it detects sensitive data in an S3 object that it analyzes to discover sensitive data. This includes analysis that Macie performs when you run a sensitive data discovery job or it performs automated sensitive data discovery.

For example, if you create and run a sensitive data discovery job and Macie detects bank account numbers in an S3 object, Macie generates a **SensitiveData:S3Object/Financial** finding for the object. Similarly, if Macie detects bank account numbers in an S3 object that it analyzes during an automated sensitive data discovery cycle, Macie generates a **SensitiveData:S3Object/Financial** finding for the object.

If Macie detects sensitive data in the same S3 object during a subsequent job run or automated sensitive data discovery cycle, Macie generates a new sensitive data finding for the object. Unlike policy findings, all sensitive data findings are treated as new (unique). Macie stores sensitive data findings for 90 days.

Macie can generate the following types of sensitive data findings for an S3 object.

SensitiveData:S3Object/Credentials

The object contains sensitive credentials data, such as AWS secret access keys or private keys.

SensitiveData:S3Object/CustomIdentifier

The object contains text that matches the detection criteria of one or more custom data identifiers. The object might contain more than one type of sensitive data.

SensitiveData:S3Object/Financial

The object contains sensitive financial information, such as bank account numbers or credit card numbers.

SensitiveData:S3Object/Multiple

The object contains more than one category of sensitive data—any combination of credentials data, financial information, personal information, or text that matches the detection criteria of one or more custom data identifiers.

SensitiveData:S3Object/Personal

The object contains sensitive personal information—personally identifiable information (PII) such as passport numbers or driver's license identification numbers, personal health

information (PHI) such as health insurance or medical identification numbers, or a combination of PII and PHI.

For information about the types of sensitive data that Macie can detect using built-in criteria and techniques, see Using managed data identifiers. For information about the types of S3 objects that Macie can analyze, see Supported storage classes and formats.

Severity scoring for Macie findings

When Amazon Macie generates a policy or sensitive data finding, it automatically assigns a severity to the finding. A finding's severity reflects the principal characteristics of the finding, which can help you assess and prioritize the finding. A finding's severity doesn't imply or otherwise indicate the criticality or importance that an affected resource might have for your organization.

For policy findings, severity is based on the nature of a potential issue with the security or privacy of an Amazon Simple Storage Service (Amazon S3) general purpose bucket. For sensitive data findings, severity is based on the nature and number of occurrences of sensitive data that Macie detected in an S3 object.

In Macie, a finding's severity is represented in two ways.

Severity level

This is a qualitative representation of severity. Severity levels range from *Low*, for least severe, to *High*, for most severe.

Severity levels appear directly on the Amazon Macie console. They're also available in JSON representations of findings on the Macie console, from the Amazon Macie API, and in sensitive data discovery results that correlate to sensitive data findings. Severity levels are also included in finding events that Macie publishes to Amazon EventBridge and findings that Macie publishes to AWS Security Hub.

Severity score

This is a numerical representation of severity. Severity scores range from *1* through *3* and map directly to severity levels:

Severity score	Severity level
1	Low

Severity score	Severity level
2	Medium
3	High

Severity scores don't appear directly on the Amazon Macie console. However, they're available in JSON representations of findings on the Macie console, from the Amazon Macie API, and in sensitive data discovery results that correlate to sensitive data findings. Severity scores are also included in finding events that Macie publishes to Amazon EventBridge. They aren't included in findings that Macie publishes to AWS Security Hub.

The topics in this section indicate how Macie determines the severity of policy findings and sensitive data findings.

Topics

Severity scoring for policy findings

The severity of a policy finding is based on the nature of a potential issue with the security or privacy of an S3 general purpose bucket. The following table lists the severity levels that Amazon Macie assigns to each type of policy finding. For a description of each type, see Types of findings.

Finding type	Severity level
Policy:IAMUser/S3BlockPublicAccessDisabled	High
Policy:IAMUser/S3BucketEncryptionDisabled	Low
Policy:IAMUser/S3BucketPublic	High
Policy:IAMUser/S3BucketReplicatedExternally	High
Policy:IAMUser/S3BucketSharedExternally	High

Finding type	Severity level
Policy:IAMUser/S3BucketSharedWithCloudFront	Medium

The severity of a policy finding doesn't change based on the number of occurrences of the finding.

Severity scoring for sensitive data findings

The severity of a sensitive data finding is based on the nature and number of occurrences of sensitive data that Amazon Macie detected in an S3 object. The following topics indicate how Macie determines the severity of each type of sensitive data finding:

- SensitiveData:S3Object/Credentials
- SensitiveData:S3Object/CustomIdentifier
- SensitiveData:S3Object/Financial
- SensitiveData:S3Object/Personal
- SensitiveData:S3Object/Multiple

For details about the types of sensitive data that Macie can detect and report in sensitive data findings, see Using managed data identifiers and Building custom data identifiers.

SensitiveData:S3Object/Credentials

A **SensitiveData:S3Object/Credentials** finding indicates that Macie detected sensitive credentials data in an S3 object. For this type of finding, Macie determines severity based on the type and number of occurrences of the credentials data that Macie detected in the object.

The following table indicates the severity levels that Macie assigns to findings that report occurrences of credentials data in an S3 object.

Sensitive data type	1 occurrence	2–99 occurrences	100 or more occurrences
AWS secret access key	High	High	High

Sensitive data type	1 occurrence	2–99 occurrences	100 or more occurrences
Google Cloud API key	High	High	High
HTTP Basic Authorization header	High	High	High
JSON Web Token (JWT)	High	High	High
OpenSSH private key	High	High	High
PGP private key	High	High	High
Public-Key Cryptography Standard (PKCS) private key	High	High	High
PuTTY private key	High	High	High
Stripe API key	High	High	High

SensitiveData:S3Object/CustomIdentifier

A **SensitiveData:S3Object/CustomIdentifier** finding indicates that an S3 object contains text that matches the detection criteria of one or more custom data identifiers. The object might contain more than one type of sensitive data.

By default, Macie assigns the **Medium** severity level to this type of finding. If the affected S3 object contains at least one occurrence of text that matches the detection criteria of at least one custom data identifier, Macie automatically assigns the **Medium** severity level to the finding. The severity of the finding doesn't change based on the number of occurrences of text that match a custom data identifier's criteria.

However, the severity of this type of finding can vary if you defined custom severity settings for a custom data identifier that produced the finding. If this is the case, Macie determines severity as follows:

- If the S3 object contains text that matches the detection criteria of only one custom data identifier, Macie determines the finding's severity based on the severity settings for that identifier.
- If the S3 object contains text that matches the detection criteria of more than one custom data identifier, Macie determines the finding's severity by evaluating the severity settings for each custom data identifier, determining which of those settings produces the highest severity, and then assigning that highest severity to the finding.

To review the severity settings for a custom data identifier, you can use the Amazon Macie console or the Amazon Macie API. To review the settings on the console, choose **Custom data identifiers** in the navigation pane, and then choose the name of the custom data identifier. The **Severity** section shows the settings. To retrieve the settings programmatically, use the GetCustomDataIdentifier operation or, if you're using the AWS Command Line Interface, run the get-custom-data-identifier command. To learn about the settings, see Configuration options for custom data identifiers.

SensitiveData:S3Object/Financial

A **SensitiveData:S3Object/Financial** finding indicates that Macie detected sensitive financial information in an S3 object. For this type of finding, Macie determines severity based on the type and number of occurrences of the financial information that Macie detected in the object.

The following table indicates the severity levels that Macie assigns to findings that report occurrences of financial information in an S3 object.

Sensitive data type	1 occurrence	2–99 occurrences	100 or more occurrences
Bank account number [1]	High	High	High
Credit card expiration date	Low	Medium	High
Credit card magnetic stripe data	High	High	High

Sensitive data type	1 occurrence	2–99 occurrences	100 or more occurrences
Credit card number [2]	High	High	High
Credit card verificat ion code	Medium	High	High

1. Severity levels are the same for any type of bank account number—a Basic Bank Account Number (BBAN), an International Bank Account Number (IBAN), or a Canadian or US bank account number.
2. Severity levels are the same for credit card numbers that are or aren't in proximity of a keyword.

If a finding reports multiple types of financial information in an S3 object, Macie determines the finding's severity by calculating the severity for each type of financial information that Macie detected, determining which type produces the highest severity, and assigning that highest severity to the finding. For example, if Macie detects 10 credit card expiration dates (**Medium** severity level) and 10 credit card numbers (**High** severity level) in an object, Macie assigns the **High** severity level to the finding.

SensitiveData:S3Object/Personal

A **SensitiveData:S3Object/Personal** finding indicates that Macie detected sensitive personal information in an S3 object. The information can be personal health information (PHI), personally identifiable information (PII), or a combination of the two. For this type of finding, Macie determines severity based on the type and number of occurrences of the personal information that Macie detected in the object.

The following table indicates the severity levels that Macie assigns to sensitive data findings that report occurrences of PHI in an S3 object.

Sensitive data type	1 occurrence	2–99 occurrences	100 or more occurrences
	High	High	High

Sensitive data type	1 occurrence	2–99 occurrences	100 or more occurrences
Drug Enforcement Agency (DEA) Registration Number			
Health Insurance Claim Number (HICN)	High	High	High
Health insurance or medical identification number	High	High	High
Healthcare Common Procedure Coding System (HCPCS) code	High	High	High
National Drug Code (NDC)	High	High	High
National Provider Identifier (NPI)	High	High	High
Unique device identifier (UDI)	Low	Medium	High

The following table indicates the severity levels that Macie assigns to sensitive data findings that report occurrences of PII in an S3 object.

Sensitive data type	1 occurrence	2–99 occurrences	100 or more occurrences
Birth date	Low	Medium	High
Driver's license identification number	Low	Medium	High

Sensitive data type	1 occurrence	2–99 occurrences	100 or more occurrences
Electoral roll number	High	High	High
Full name	Low	Medium	High
Global Positioning System (GPS) coordinates	Low	Medium	Medium
HTTP cookie	Low	Medium	High
Mailing address	Low	Medium	High
National identification number	High	High	High
National Insurance Number (NINO)	High	High	High
Passport number	Medium	High	High
Permanent residence number	High	High	High
Phone number	Low	Medium	High
Public transportation card number	Medium	Medium	High
Social Insurance Number (SIN)	High	High	High
Social Security number (SSN)	High	High	High

Sensitive data type	1 occurrence	2–99 occurrences	100 or more occurrences
Taxpayer identification or reference number *	High	High	High
Vehicle identification number (VIN)	Low	Low	Medium

* Exceptions are: CUIT numbers for organizations in Argentina (`ARGENTINA_ORGANIZATION_TAX_IDENTIFICATION_NUMBER`), NIT numbers for organizations in Colombia (`COLOMBIA_ORGANIZATION_NIT_NUMBER`), and RFC numbers for organizations in Mexico (`MEXICO_ORGANIZATION_RFC_NUMBER`). For those types, the severity levels are: **Medium** for 1–99 occurrences, and **High** for 100 or more occurrences.

If a finding reports multiple types of PHI, PII, or both PHI and PII in an object, Macie determines the finding's severity by calculating the severity for each type, determining which type produces the highest severity, and assigning that highest severity to the finding.

For example, if Macie detects 10 full names (**Medium** severity level) and 5 passport numbers (**High** severity level) in an object, Macie assigns the **High** severity level to the finding. Similarly, if Macie detects 10 full names (**Medium** severity level) and 10 health insurance identification numbers (**High** severity level) in an object, Macie assigns the **High** severity level to the finding.

SensitiveData:S3Object/Multiple

A **SensitiveData:S3Object/Multiple** finding indicates that Macie detected multiple categories of sensitive data in an S3 object. The sensitive data can be any combination of credentials data, financial information, personal information, or text that matches the detection criteria of one or more custom data identifiers.

For this type of finding, Macie determines severity by calculating the severity for each type of sensitive data that Macie detected (as indicated in the preceding topics), determining which type produces the highest severity, and assigning that highest severity to the finding.

For example, if Macie detects 10 full names (**Medium** severity level) and 10 AWS secret access keys (**High** severity level) in an object, Macie assigns the **High** severity level to the finding.

Working with Macie sample findings

To explore and learn about the different types of findings that Amazon Macie can generate, you can create sample findings. Sample findings use example data and placeholder values to demonstrate the kinds of information that each type of finding might contain.

For example, the **Policy:IAMUser/S3BucketPublic** sample finding contains details about a fictitious Amazon Simple Storage Service (Amazon S3) bucket. The finding's details include example data about an actor and action that changed the access control list (ACL) for the bucket and made the bucket publicly accessible. Similarly, the **SensitiveData:S3Object/Multiple** sample finding contains details about a fictitious Microsoft Excel workbook. The finding's details include example data about the types and location of sensitive data in the workbook.

In addition to familiarizing yourself with the information that different types of findings might contain, you can use sample findings to test integration with other applications, services, and systems. Depending on the suppression rules for your account, Macie can publish sample findings to Amazon EventBridge as events. The example data in these events can help you develop and test automated solutions for monitoring and processing findings with EventBridge. Depending on the publication settings for your account, Macie can also publish sample findings to AWS Security Hub. This means that you can also use sample findings to develop and test solutions for evaluating Macie findings with Security Hub. For information about publishing findings to these services, see Monitoring and processing findings.

Topics

Creating sample findings

You can create sample findings by using the Amazon Macie console or the Amazon Macie API. If you use the console, Macie automatically generates one sample finding for each type of finding that Macie supports. If you use the API, you can create a sample for each type, or only certain types that you specify.

Console

Follow these steps to create sample findings by using the Amazon Macie console.

To create sample findings

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Settings**.
3. Under **Sample findings**, choose **Generate sample findings**.

API

To create sample findings programmatically, use the CreateSampleFindings operation of the Amazon Macie API. When you submit your request, optionally use the `findingTypes` parameter to specify only certain types of sample findings to create. To automatically create samples of all types, don't include this parameter in your request.

To create sample findings by using the AWS Command Line Interface (AWS CLI), run the create-sample-findings command. To automatically create samples of all types of findings, don't include the `finding-types` parameter. To create samples of only certain types of findings, include this parameter and specify the types of sample findings to create. For example:

```
C:\> aws macie2 create-sample-findings --finding-types "SensitiveData:S3Object/
Multiple" "Policy:IAMUser/S3BucketPublic"
```

Where *`SensitiveData:S3Object/Multiple`* is a type of sensitive data finding to create and *`Policy:IAMUser/S3BucketPublic`* is a type of policy finding to create.

If the command runs successfully, Macie returns an empty response.

If you create sample findings again within 90 days, Macie generates a new finding for each type of sensitive data finding that you create. For policy findings, Macie updates each existing sample finding by incrementing the count of occurrences and updating details about when the subsequent occurrence occurred.

Reviewing sample findings

To help you identify sample findings, Amazon Macie sets the value for the **Sample** field of each sample finding to *True*. In addition, the name of the affected S3 bucket is the same for all sample

findings: *macie-sample-finding-bucket*. If you review sample findings by using **Findings** pages on the Amazon Macie console, Macie also displays the **[SAMPLE]** prefix in the **Finding type** field for each sample finding.

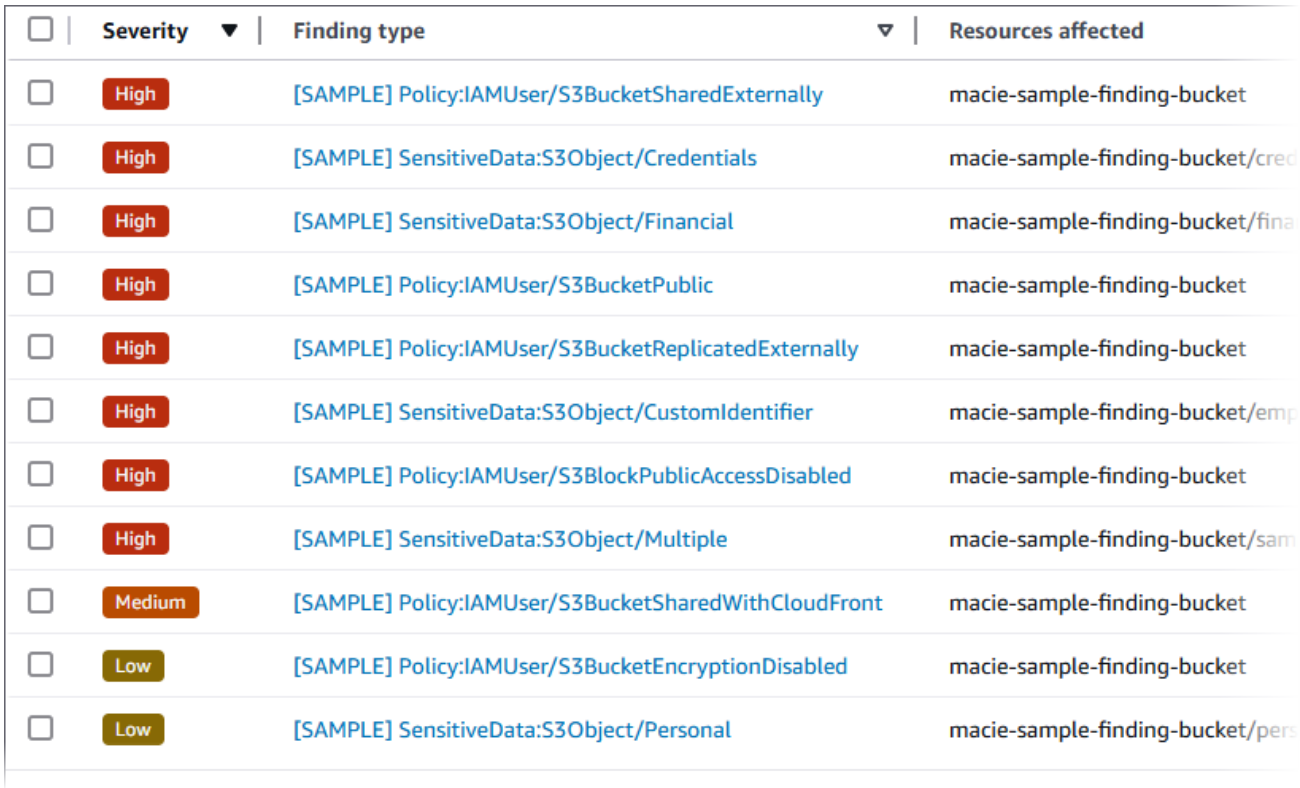
Console

Follow these steps to review sample findings by using the Amazon Macie console.

To review sample findings

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Findings**.
3. On the **Findings** page, do any of the following:

   - In the **Finding type** column, locate findings whose type begins with **[SAMPLE]**, as shown in the following image.



   - By using the **Filter criteria** box above the table, filter the table to display only sample findings. To do this, place your cursor in the box. In the list of fields that appears, choose **Sample**. Then choose **True**, and then choose **Apply**.
4. To review the details of a specific sample finding, choose the finding. The details panel displays information for the finding.

You can also download and save the details of one or more sample findings as a JSON file. To do this, select the checkbox for each sample finding that you want to download and save. Then choose **Export (JSON)** on the **Actions** menu at the top of the **Findings** page. In the window that appears, choose **Download**. For detailed descriptions of the JSON fields that a finding can include, see Findings in the *Amazon Macie API Reference*.

API

To review sample findings programmatically, first use the ListFindings operation of the Amazon Macie API to retrieve the unique identifier (`findingId`) for each sample finding that you created. Then use the GetFindings operation to retrieve the details of those findings.

When you submit the **ListFindings** request, you can specify filter criteria to include only sample findings in the results. To do this, add a filter condition where the value for the `sample` field is `true`. If you're using the AWS CLI, run the list-findings command and use the `finding-criteria` parameter to specify the filter condition. For example:

```
C:\> aws macie2 list-findings --finding-criteria={\"criterion\":{\"sample\":{\"eq\":
[\"true\"]}}}
```

If your request succeeds, Macie returns a `findingIds` array. The array lists the unique identifier for each sample finding for your account in the current AWS Region.

To then retrieve the details of the sample findings, specify these unique identifiers in a **GetFindings** request or, for the AWS CLI, when you run the get-findings command.

Suppressing sample findings

Like other findings, Amazon Macie stores sample findings for 90 days. After you finish reviewing and experimenting with the samples, you can optionally archive them by creating a suppression rule. If you do this, the sample findings stop appearing by default on the console and their status changes to *archived*.

To archive sample findings by using the Amazon Macie console, configure the rule to archive findings where the value for the **Sample** field is **True**. To archive sample findings by using the Amazon Macie API, configure the rule to archive findings where the value for the `sample` field is `true`.

Reviewing Macie findings by using the console

Amazon Macie monitors your AWS environment and generates policy findings when it detects potential policy violations or issues with the security or privacy of your Amazon Simple Storage Service (Amazon S3) general purpose buckets. Macie generates sensitive data findings when it detects sensitive data in S3 objects. Macie stores your policy and sensitive data findings for 90 days.

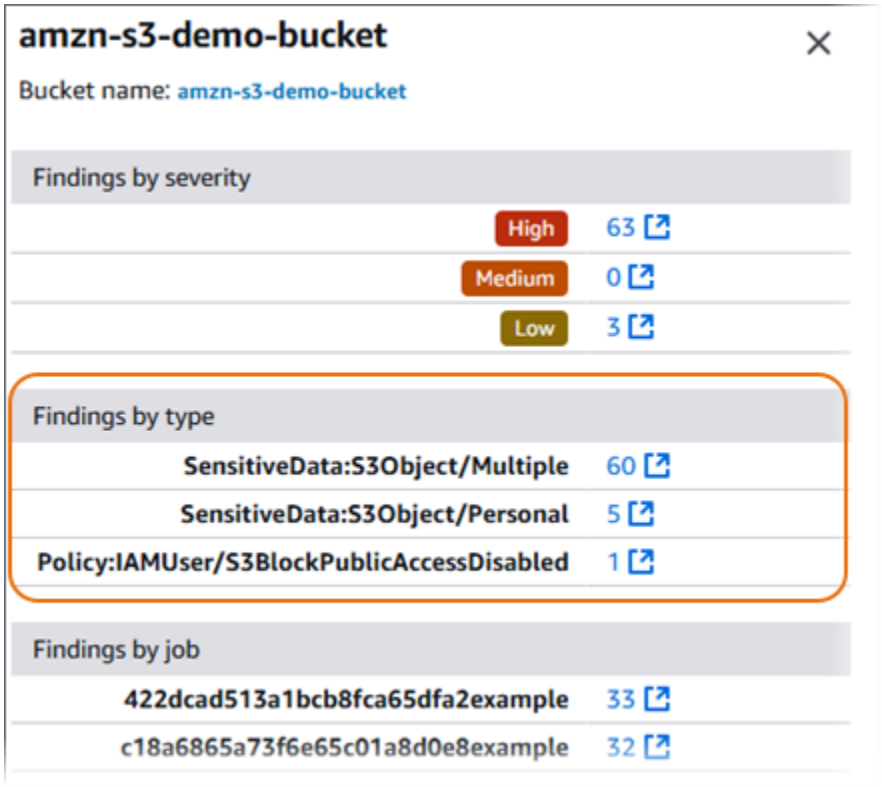
Each finding specifies a finding type and severity rating. Additional details include information about the affected resource and when and how Macie found the issue or sensitive data reported by the finding. The severity and details of each finding vary depending on the type and nature of the finding.

By using the Amazon Macie console, you can review and analyze findings, and access the details of individual findings. You can also export one or more findings to a JSON file. To streamline your analysis, the console offers several options that can help you build custom views of findings.

Use predefined groupings

Use specific pages to review findings that are grouped by criteria such as affected S3 bucket, finding type, or sensitive data discovery job. With these pages, you can review aggregated statistics for each group, such as the count of findings by severity. You can also drill down to review the details of individual findings in a group, and you can apply filters to refine your analysis.

For example, if you group all findings by S3 bucket and note that a particular bucket has a policy violation, you can quickly determine whether there are also sensitive data findings for the bucket. To do this, choose **By bucket** in the navigation pane (under **Findings**), and then choose the bucket. In the details panel that appears, the **Findings by type** section lists the types of findings that apply to the bucket, as shown in the following image.

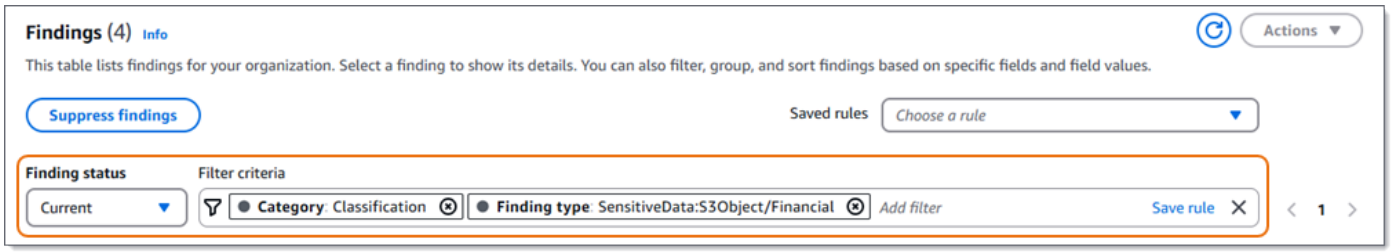


To investigate a specific type, choose the number for the type. Macie displays a table of all the findings that match the selected type and apply to the S3 bucket. To refine the results, filter the table.

Create and apply filters

Use specific finding attributes to include or exclude certain findings from a **Findings** table. A *finding attribute* is a field that stores specific data for a finding, such as finding type, severity, or the name of the affected S3 bucket. If you filter a table, you can more easily identify findings that have specific characteristics. Then you can drill down to review the details of those findings.

For example, to review all of your sensitive data findings, add filter criteria for the **Category** field. To refine the results and include only a specific type of sensitive data finding, add filter criteria for the **Finding type** field. For example:



To then review the details of a particular finding, choose the finding. The details panel displays information for the finding.

You can also sort findings in ascending or descending order by certain fields. To do this, choose the column heading for the field. To change the sort order, choose the column heading again.

To review findings by using the console

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Findings**. The **Findings** page displays findings that Macie created or updated for your account in the current AWS Region during the past 90 days. By default, this doesn't include findings that were suppressed by a suppression rule.
3. To pivot on and review findings by a predefined logical group, choose **By bucket**, **By type**, or **By job** in the navigation pane (under **Findings**). Then choose an item in the table. In the details panel, choose the link for the field to pivot on.
4. To filter the findings by specific criteria, use the filter options above the table:

   - To display findings that were suppressed by a suppression rule, use the **Finding status** menu. Choose **All** to display both suppressed and unsuppressed findings, or choose **Archived** to display only suppressed findings. To then hide suppressed findings again, choose **Current**.
   - To display only those findings that have a specific attribute, use the **Filter criteria** box. Place your cursor in the box and add a filter condition for the attribute. To further refine the results, add conditions for additional attributes. To then remove a condition, choose the remove condition icon () for the condition to remove.

   For more information about filtering findings, see Creating and applying filters to Macie findings.
5. To sort the findings by a specific field, choose the column heading for the field. To change the sort order, choose the column heading again.
6. To review the details of a specific finding, choose the finding. The details panel displays information for the finding.

Tips

In the details panel, you can pivot and drill down on certain fields. To show findings that have the same value for a field, choose  in the field. Choose  to show findings that have other values for the field.

For a sensitive data finding, you can also use the details panel to investigate sensitive data that Macie found in the affected S3 object:

- To locate occurrences of a specific type of sensitive data, choose the numeric link in the field for that type of data. Macie displays information (in JSON format) about where Macie found the data. For more information, see Locating sensitive data.
- To retrieve samples of the sensitive data that Macie found, choose **Review** in the **Reveal samples** field. For more information, see Retrieving sensitive data samples.
- To navigate to the corresponding sensitive data discovery result, choose the link in the **Detailed result location** field. Macie opens the Amazon S3 console and displays the file or folder that contains the discovery result. For more information, see Storing and retaining sensitive data discovery results.

7. To download and save the details of one or more findings as a JSON file, select the checkbox for each finding to download and save. Then choose **Export (JSON)** on the **Actions** menu. In the window that appears, choose **Download**. For detailed descriptions of the JSON fields that a finding can include, see Findings in the *Amazon Macie API Reference*.

In certain cases, a finding might not include all the details of an affected S3 bucket. This can occur if you store more than 10,000 buckets in Amazon S3. Macie maintains complete inventory data for only 10,000 buckets for an account—the 10,000 buckets that were most recently created or changed. To review additional details for an affected bucket, you can use data in the finding to determine the bucket's name, the account ID for the AWS account that owns the bucket, and the AWS Region that stores the bucket. You can then use Amazon S3 to review all the bucket's details.

Filtering Macie findings

To perform targeted analysis and to analyze findings more efficiently, you can filter Amazon Macie findings. With filters, you build custom views and queries for findings, which can help you identify and focus on findings that have specific characteristics. Use the Amazon Macie console to filter findings, or submit queries programmatically using the Amazon Macie API.

When you create a filter, you use specific attributes of findings to define criteria for including or excluding findings from a view or from query results. A *finding attribute* is a field that stores specific data for a finding, such as severity, type, or the name of the S3 bucket that a finding applies to.

In Macie, a filter consists of one or more conditions. Each condition, also referred to as a *criterion*, consists of three parts:

- An attribute-based field, such as **Severity** or **Finding type**.
- An operator, such as *equals* or *not equals*.
- One or more values. The type and number of values depends on the field and operator that you choose.

If you create a filter that you want to use again, you can save it as a *filter rule*. A *filter rule* is a set of filter criteria that you create and save to reapply when you review findings on the Amazon Macie console.

You can also save a filter as a *suppression rule*. A *suppression rule* is a set of filter criteria that you create and save to automatically archive findings that match the criteria of the rule. To learn about suppression rules, see Suppressing findings.

Topics

Fundamentals of filtering Macie findings

When you filter findings, keep the following features and guidelines in mind. Also note that filtered results are limited to the preceding 90 days and the current AWS Region. Amazon Macie stores your findings for only 90 days in each AWS Region.

Topics

Using multiple conditions in a filter

A filter can include one or more conditions. Each condition, also referred to as a *criterion*, consists of three parts:

- An attribute-based field, such as **Severity** or **Finding type**. For a list of fields that you can use, see Fields for filtering Macie findings.
- An operator, such as *equals* or *not equals*. For a list of operators that you can use, see Using operators in conditions.
- One or more values. The type and number of values depends on the field and operator that you choose.

If a filter contains multiple conditions, Amazon Macie uses AND logic to join the conditions and evaluate the filter criteria. This means that a finding matches the filter criteria only if it matches *all* the conditions in the filter.

For example, if you add a condition to include only high-severity findings and add another condition to include only sensitive data findings, Macie returns all high-severity, sensitive data findings. In other words, Macie excludes all policy findings and all medium-severity and low-severity sensitive data findings.

You can use a field only once in a filter. However, you can specify multiple values for many fields.

For example, if a condition uses the **Severity** field to include only high-severity findings, you can't use the **Severity** field in another condition to include medium-severity or low-severity findings.

Instead, specify multiple values for the existing condition, or use a different operator for the existing condition. For example, to include all medium-severity and high-severity findings, add a **Severity** *equals Medium, High* condition or add a **Severity** *not equals Low* condition.

Specifying values for fields

When you specify a value for a field, the value has to conform to the underlying data type for the field. Depending on the field, you can specify one of the following types of values.

Array of text (strings)

Specifies a list of text (string) values for a field. Each string correlates to a predefined or existing value for a field—for example, *High* for the **Severity** field, *SensitiveData:S3Object/Financial* for the **Finding type** field, or the name of an S3 bucket for the **S3 bucket name** field.

If you use an array, note the following:

- Values are case sensitive.
- You can't specify partial values or use wildcard characters in values. You have to specify a complete, valid value for the field.

For example, to filter findings for an S3 bucket named *my-S3-bucket*, enter **`my-S3-bucket`** as the value for the **S3 bucket name** field. If you enter any other value, such as **`my-s3-bucket`** or **`my-S3`**, Macie won't return findings for the bucket.

For a list of valid values for each field, see Fields for filtering Macie findings.

You can specify as many as 50 values in an array. How you specify the values depends on whether you use the Amazon Macie console or the Amazon Macie API, as discussed in Specifying multiple values for a field.

Boolean

Specifies one of two mutually exclusive values for a field.

If you use the Amazon Macie console to specify this type of value, the console provides a list of values to choose from. If you use the Amazon Macie API, specify `true` or `false` for the value.

Date/Time (and time ranges)

Specifies an absolute date and time for a field. If you specify this type of value, you have to specify both a date and time.

On the Amazon Macie console, date and time values are in your local time zone and use 24-hour notation. In all other contexts, these values are in Coordinated Universal Time (UTC) and extended ISO 8601 format—for example `2020-09-01T14:31:13Z` for 2:31:13 PM UTC September 1, 2020.

If a field stores a date/time value, you can use the field to define a fixed or relative time range. For example, you can include only those findings that were created between two specific dates and times, or only those findings that were created before or after a specific date and time. How you define a time range depends on whether you use the Amazon Macie console or the Amazon Macie API:

- On the console, use a date picker or enter text directly in the **From** and **To** boxes.
- With the API, define a fixed time range by adding a condition that specifies the first date and time in the range, and add another condition that specifies the last date and time in the range. If you do this, Macie uses AND logic to join the conditions. To define a relative time range, add one condition that specifies the first or last date and time in the range. Specify the values as Unix timestamps in milliseconds—for example, `1604616572653` for 22:49:32 UTC November 5, 2020.

On the console, time ranges are inclusive. With the API, time ranges can be inclusive or exclusive, depending on the operator that you choose.

Number (and numeric ranges)

Specifies a long integer for a field.

If a field stores a numeric value, you can use the field to define a fixed or relative numeric range. For example, you can include only those findings that report 50-90 occurrences of sensitive data in an S3 object. How you define a numeric range depends on whether you use the Amazon Macie console or the Amazon Macie API:

- On the console, use the **From** and **To** boxes to enter the lowest and highest numbers in the range, respectively.
- With the API, define a fixed numeric range by adding a condition that specifies the lowest number in the range, and add another condition that specifies the highest number in the range. If you do this, Macie uses AND logic to join the conditions. To define a relative numeric range, add one condition that specifies the lowest or highest number in the range.

On the console, numeric ranges are inclusive. With the API, numeric ranges can be inclusive or exclusive, depending on the operator that you choose.

Text (string)

Specifies a single text (string) value for a field. The string correlates to a predefined or existing value for a field—for example, *High* for the **Severity** field, the name of an S3 bucket for the **S3 bucket name** field, or the unique identifier for a sensitive data discovery job for the **Job ID** field.

If you specify a single text string, note the following:

- Values are case sensitive.
- You can't use partial values or use wildcard characters in values. You have to specify a complete, valid value for the field.

For example, to filter findings for an S3 bucket named *my-S3-bucket*, enter **`my-S3-bucket`** as the value for the **S3 bucket name** field. If you enter any other value, such as **`my-s3-bucket`** or **`my-S3`**, Macie won't return findings for the bucket.

For a list of valid values for each field, see Fields for filtering Macie findings.

Specifying multiple values for a field

With certain fields and operators, you can specify multiple values for a field. If you do this, Amazon Macie uses OR logic to join the values and evaluate the filter criteria. This means that a finding matches the criteria if it has *any* of the values for the field.

For example, if you add a condition to include findings where the value for the **Finding type** field equals *SensitiveData:S3Object/Financial, SensitiveData:S3Object/Personal*, Macie returns sensitive data findings for S3 objects that contain only financial information, and S3 objects that contain only personal information. In other words, Macie excludes all policy findings. Macie also excludes all sensitive data findings for objects that contain other types of sensitive data or multiple types of sensitive data.

The exception is conditions that use the *eqExactMatch* operator. For this operator, Macie uses AND logic to join the values and evaluate the filter criteria. This means that a finding matches the criteria only if it has *all* the values for the field and *only* those values for the field. To learn more about this operator, see Using operators in conditions.

How you specify multiple values for a field depends on whether you use the Amazon Macie API or the Amazon Macie console. With the API, you use an array that lists the values.

On the console, you typically choose the values from a list. However, for some fields, you have to add a distinct condition for each value. For example, to include findings for data that Macie detected using certain custom data identifiers, do the following:

1. Place your cursor in the **Filter criteria** box and then choose the **Custom data identifier name** field. Enter the name of a custom data identifier, and then choose **Apply**.
2. Repeat the preceding step for each additional custom data identifier that you want to specify for the filter.

For a list of fields that you need to do this for, see Fields for filtering Macie findings.

Using operators in conditions

You can use the following types of operators in individual conditions.

Equals (eq)

Matches (=) any value specified for the field. You can use the *equals* operator with the following types of values: array of text (strings), Boolean, date/time, number, and text (string).

For many fields, you can use this operator and specify as many as 50 values for the field. If you do this, Amazon Macie uses OR logic to join the values. This means that a finding matches the criteria if it has *any* of the values specified for the field.

For example:

- To include findings that report occurrences of financial information, personal information, or both financial and personal information, add a condition that uses the **Sensitive data category** field and this operator, and specify *Financial information* and *Personal information* as the values for the field.
- To include findings that report occurrences of credit card numbers, mailing addresses, or both credit card numbers and mailing addresses, add a condition for the **Sensitive data detection type** field, use this operator, and specify *CREDIT_CARD_NUMBER* and *ADDRESS* as the values for the field.

If you use the Amazon Macie API to define a condition that uses this operator with a date/time value, specify the value as a Unix timestamp in milliseconds—for example, 1604616572653 for 22:49:32 UTC November 5, 2020.

Equals exact match (eqExactMatch)

Exclusively matches all the values specified for the field. You can use the *equals exact match* operator with a select set of fields.

If you use this operator and specify multiple values for a field, Macie uses AND logic to join the values. This means that a finding matches the criteria only if it has *all* the values specified for the field and *only* those values for the field. You can specify as many as 50 values for the field.

For example:

- To include findings that report occurrences of credit card numbers and no other type of sensitive data, add a condition for the **Sensitive data detection type** field, use this operator, and specify *CREDIT_CARD_NUMBER* as the only value for the field.
- To include findings that report occurrences of both credit card numbers and mailing addresses (and no other types of sensitive data), add a condition for the **Sensitive data detection type** field, use this operator, and specify *CREDIT_CARD_NUMBER* and *ADDRESS* as the values for the field.

Because Macie uses AND logic to join the values for a field, you can't use this operator in combination with any other operators for the same field. In other words, if you use the *equals exact match* operator with a field in one condition, you have to use it in all other conditions that use the same field.

Like other operators, you can use the *equals exact match* operator in more than one condition in a filter. If you do this, Macie uses AND logic to join the conditions and evaluate the filter. This means that a finding matches the filter criteria only if it has *all* the values specified by *all* the conditions in the filter.

For example, to include findings that were created after a certain time, report occurrences of credit card numbers, and don't report any other type of sensitive data, do the following:

1. Add a condition that uses the **Created at** field, uses the *greater than* operator, and specifies the starting date and time for the filter.
2. Add another condition that uses the **Sensitive data detection type** field, uses the *equals exact match* operator, and specifies *CREDIT_CARD_NUMBER* as the only value for the field.

You can use the *equals exact match* operator with the following fields:

- Custom data identifier ID (`customDataIdentifiers.detections.arn`)

- Custom data identifier name (`customDataIdentifiers.detections.name`)
- S3 bucket tag key (`resourcesAffected.s3Bucket.tags.key`)
- S3 bucket tag value (`resourcesAffected.s3Bucket.tags.value`)
- S3 object tag key (`resourcesAffected.s3Object.tags.key`)
- S3 object tag value (`resourcesAffected.s3Object.tags.value`)
- Sensitive data detection type (`sensitiveData.detections.type`)
- Sensitive data category (`sensitiveData.category`)

In the preceding list, the parenthetical name uses dot notation to indicate the name of the field in JSON representations of findings and the Amazon Macie API.

Greater than (gt)

Is greater than (>) the value specified for the field. You can use the *greater than* operator with number and date/time values.

For example, to include only those findings that report more than 90 occurrences of sensitive data in an S3 object, add a condition that uses the **Sensitive data total count** field and this operator, and specify 90 as the value for the field. To do this on the Amazon Macie console, enter **91** in the **From** box, don't enter a value in the **To** box, and then choose **Apply**. Numeric and time-based comparisons are inclusive on the console.

If you use the Amazon Macie API to define a time range that uses this operator, you have to specify the date/time values as Unix timestamps in milliseconds—for example, 1604616572653 for 22:49:32 UTC November 5, 2020.

Greater than or equal to (gte)

Is greater than or equal to (>=) the value specified for the field. You can use the *greater than or equal to* operator with number and date/time values.

For example, to include only those findings that report 90 or more occurrences of sensitive data in an S3 object, add a condition that uses the **Sensitive data total count** field and this operator, and specify 90 as the value for the field. To do this on the Amazon Macie console, enter **90** in the **From** box, don't enter a value in the **To** box, and then choose **Apply**.

If you use the Amazon Macie API to define a time range that uses this operator, you have to specify the date/time values as Unix timestamps in milliseconds—for example, 1604616572653 for 22:49:32 UTC November 5, 2020.

Less than (lt)

Is less than (<) the value specified for the field. You can use the *less than* operator with number and date/time values.

For example, to include only those findings that report fewer than 90 occurrences of sensitive data in an S3 object, add a condition that uses the **Sensitive data total count** field and this operator, and specify 90 as the value for the field. To do this on the Amazon Macie console, enter **89** in the **To** box, don't enter a value in the **From** box, and then choose **Apply**. Numeric and time-based comparisons are inclusive on the console.

If you use the Amazon Macie API to define a time range that uses this operator, you have to specify the date/time values as Unix timestamps in milliseconds—for example, 1604616572653 for 22:49:32 UTC November 5, 2020.

Less than or equal to (lte)

Is less than or equal to (<=) the value specified for the field. You can use the *less than or equal to* operator with number and date/time values.

For example, to include only those findings that report 90 or fewer occurrences of sensitive data in an S3 object, add a condition that uses the **Sensitive data total count** field and this operator, and specify 90 as the value for the field. To do this on the Amazon Macie console, enter **90** in the **To** box, don't enter a value in the **From** box, and then choose **Apply**.

If you use the Amazon Macie API to define a time range that uses this operator, you have to specify the date/time values as Unix timestamps in milliseconds—for example, 1604616572653 for 22:49:32 UTC November 5, 2020.

Not equals (neq)

Doesn't match (≠) any value specified for the field. You can use the *not equals* operator with the following types of values: array of text (strings), Boolean, date/time, number, and text (string).

For many fields, you can use this operator and specify as many as 50 values for the field. If you do this, Macie uses OR logic to join the values. This means that a finding matches the criteria if it doesn't have *any* of the values specified for the field.

For example:

- To exclude findings that report occurrences of financial information, personal information, or both financial and personal information, add a condition that uses the **Sensitive data**

category field and this operator, and specify *Financial information* and *Personal information* as the values for the field.

- To exclude findings that report occurrences of credit card numbers, add a condition for the **Sensitive data detection type** field, use this operator, and specify *CREDIT_CARD_NUMBER* as the value for the field.
- To exclude findings that report occurrences of credit card numbers, mailing addresses, or both credit card numbers and mailing addresses, add a condition for the **Sensitive data detection type** field, use this operator, and specify *CREDIT_CARD_NUMBER* and *ADDRESS* as the values for the field.

If you use the Amazon Macie API to define a condition that uses this operator with a date/time value, specify the value as a Unix timestamp in milliseconds—for example, `1604616572653` for 22:49:32 UTC November 5, 2020.

Fields for filtering Macie findings

To help you analyze findings more efficiently, the Amazon Macie console and the Amazon Macie API provide access to several sets of fields for filtering findings:

- **Common fields** – These fields store data that applies to any type of finding. They correlate to common attributes of findings, such as severity, finding type, and finding ID.
- **Affected resource fields** – These fields store data about the resources that a finding applies to, such as the name, tags, and encryption settings for an affected S3 bucket or object.
- **Fields for policy findings** – These fields store data that's specific to policy findings, such as the action that produced a finding, and the entity that performed the action.
- **Fields for sensitive data findings** – These fields store data that's specific to sensitive data findings, such as the category and types of sensitive data that Macie found in an affected S3 object.

A filter can use a combination of fields from any of the preceding sets. The topics in this section list and describe individual fields in each set. For additional details about these fields, including any relationships between the fields, see Findings in the *Amazon Macie API Reference*.

Topics

- Fields for policy findings
- Fields for sensitive data findings

Common fields

The following table lists and describes fields that you can use to filter findings based on common finding attributes. These fields store data that applies to any type of finding.

In the table, the **Field** column indicates the name of the field on the Amazon Macie console. The **JSON field** column uses dot notation to indicate the name of the field in JSON representations of findings and the Amazon Macie API. The **Description** column provides a brief description of the data that the field stores, and indicates any requirements for filter values. The table is sorted in ascending alphabetical order by field, and then by JSON field.

Field	JSON field	Description
Account ID*	`accountId`	The unique identifier for the AWS account that the finding applies to. This is typically the account that owns the affected resource.
—	`archived`	A Boolean value that specifies whether the finding was suppressed (automatically archived) by a suppression rule. To use this field in a filter on the console, choose an option on the **Finding status** menu: **Archived** (suppressed only), **Current** (unsuppressed only), or **All** (both suppressed and unsuppressed).
Category	`category`	The category of the finding.

Field	JSON field	Description
		The console provides a list of values to choose from when you add this field to a filter. In the API, valid values are: `CLASSIFICATION` , for a sensitive data finding; and, `POLICY`, for a policy finding.
—	`count`	The total number of occurrences of the finding. For sensitive data findings, this value is always 1. All sensitive data findings are considered unique. This field isn't available as a filter option on the console. With the API, you can use this field to define a numeric range for a filter.
Created at	`createdAt`	The date and time when Macie created the finding. You can use this field to define a time range for a filter.
Finding ID*	`id`	The unique identifier for the finding. This is a random string that Macie generates and assigns to a finding when it creates the finding.

Field	JSON field	Description
Finding type*	`type`	The type of the finding—for example, `SensitiveData:S3Object/Personal` or `Policy:IAMUser/S3BucketPublic`. The console provides a list of values to choose from when you add this field to a filter. For a list of valid values in the API, see FindingType in the *Amazon Macie API Reference*.
Region	`region`	The AWS Region that Macie created the finding in—for example, `us-east-1` or `ca-central-1`.
Sample	`sample`	A Boolean value that specifies whether the finding is a sample finding. A *sample finding* is a finding that uses example data and placeholder values to demonstrate what a finding might contain. The console provides a list of values to choose from when you add this field to a filter.

Field	JSON field	Description
Severity	`severity.description`	The qualitative representation of the finding's severity. The console provides a list of values to choose from when you add this field to a filter. In the API, valid values are: `Low`, `Medium`, and `High`.
Updated at	`updatedAt`	The date and time when the finding was last updated. For sensitive data findings, this value is the same as the value for the **Created at** field. All sensitive data findings are considered new (unique). You can use this field to define a time range for a filter.

* To specify multiple values for this field on the console, add a condition that uses the field and specifies a distinct value for the filter, and then repeat that step for each additional value. To do this with the API, use an array that lists the values to use for the filter.

Affected resource fields

The following tables list and describe the fields that you can use to filter findings based on the type of resource that a finding applies to: an S3 bucket or an S3 object.

S3 bucket

This table lists and describes fields that you can use to filter findings based on characteristics of the S3 bucket that a finding applies to.

In the table, the **Field** column indicates the name of the field on the Amazon Macie console. The **JSON field** column uses dot notation to indicate the name of the field in JSON representations of

findings and the Amazon Macie API. (Longer JSON field names use the newline character sequence (\n) to improve readability.) The **Description** column provides a brief description of the data that the field stores, and indicates any requirements for filter values. The table is sorted in ascending alphabetical order by field, and then by JSON field.

Field	JSON field	Description
—	`resourcesAffected.s3Bucket.createdAt`	The date and time when the affected bucket was created, or changes such as edits to the bucket's policy were most recently made to the affected bucket. This field isn't available as a filter option on the console. With the API, you can use this field to define a time range for a filter.

S3 bucket default encryption	`resourcesAffected.s3Bucket.\n`
	
	`defaultServerSideEncryption.encryptionType`
		The server-side encryption algorithm that's used by default to encrypt objects that are added to the affected bucket.
		
		The console provides a list of values to choose from when you add this field to a filter. For a list of valid values for the API, see EncryptionType in the *Amazon Macie API Reference*.

S3 bucket encryption KMS key id*	`resourcesAffected.s3Bucket.\n`	The Amazon Resource Name (ARN) or unique identifier

Field	JSON field	Description
	`defaultServerSideEncryption.kmsMasterKeyId`	(key ID) for the AWS KMS key that's used by default to encrypt objects that are added to the affected bucket.
S3 bucket encryption required by bucket policy	`resourcesAffected.s3Bucket.allowsUnencryptedObjectUploads`	Specifies whether the bucket policy for the affected bucket requires server-side encryption of objects when objects are added to the bucket. The console provides a list of values to choose from when you add this field to a filter. For a list of valid values for the API, see S3Bucket in the *Amazon Macie API Reference*.
S3 bucket name*	`resourcesAffected.s3Bucket.name`	The full name of the affected bucket.
S3 bucket owner display name*	`resourcesAffected.s3Bucket.owner.displayName`	The display name of the AWS user who owns the affected bucket.

Field	JSON field	Description
S3 bucket public access permission	`resourcesAffected.s3Bucket.publicAccess.effectivePermission`	Specifies whether the affected bucket is publicly accessible based on a combination of permissions settings that apply to the bucket. The console provides a list of values to choose from when you add this field to a filter. For a list of valid values for the API, see BucketPublicAccess in the *Amazon Macie API Reference*.

—	`resourcesAffected.s3Bucket.publicAccess.\n`
	
	`permissionConfiguration.accountLevelPermissions.\n`
	
	`blockPublicAccess.blockPublicAcls`
		A Boolean value that specifies whether Amazon S3 blocks public access control lists (ACLs) for the affected bucket and objects in the bucket. This is an account-level, block public access setting for the bucket.
		
		This field isn't available as a filter option on the console.

Field	JSON field	Description

—	`resourcesAffected.s3Bucket.publicAccess.\n`
	
	`permissionConfiguration.accountLevelPermissions.\n`
	
	`blockPublicAccess.blockPublicPolicy`
		A Boolean value that specifies whether Amazon S3 blocks public bucket policies for the affected bucket. This is an account-level, block public access setting for the bucket.
		
		This field isn't available as a filter option on the console.

—	`resourcesAffected.s3Bucket.publicAccess.\n`
	
	`permissionConfiguration.accountLevelPermissions.\n`
	
	`blockPublicAccess.ignorePublicAcls`
		A Boolean value that specifies whether Amazon S3 ignores public ACLs for the affected bucket and objects in the bucket. This is an account-level, block public access setting for the bucket.
		
		This field isn't available as a filter option on the console.

—	`resourcesAffected.s3Bucket.publicAccess.\n`
	
	`permissionConfiguration.accountLevelPermissions.\n`
	
	`blockPublicAccess.restrictPublicBuckets`
		A Boolean value that specifies whether Amazon S3 restricts public bucket policies for the affected bucket. This is an account-level, block public access setting for the bucket.
		
		This field isn't available as a filter option on the console.

Field	JSON field	Description

—	`resourcesAffected.s3Bucket.publicAccess.\n`
	
	`permissionConfiguration.bucketLevelPermissions.\n`
	
	`accessControlList.allowsPublicReadAccess`
		A Boolean value that specifies whether the bucket-level ACL for the affected bucket grants the general public with read access permissions for the bucket.
		
		This field isn't available as a filter option on the console.

—	`resourcesAffected.s3Bucket.publicAccess.\n`
	
	`permissionConfiguration.bucketLevelPermissions.\n`
	
	`accessControlList.allowsPublicWriteAccess`
		A Boolean value that specifies whether the bucket-level ACL for the affected bucket grants the general public with write access permissions for the bucket.
		
		This field isn't available as a filter option on the console.

—	`resourcesAffected.s3Bucket.publicAccess.\n`
	
	`permissionConfiguration.bucketLevelPermissions.\n`
	
	`blockPublicAccess.blockPublicAcls`
		A Boolean value that specifies whether Amazon S3 blocks public ACLs for the affected bucket and objects in the bucket. This is a bucket-level, block public access setting for a bucket.
		
		This field isn't available as a filter option on the console.

Field	JSON field	Description

—	`resourcesAffected.s3Bucket.publicAccess.\n`
	
	`permissionConfiguration.bucketLevelPermissions.\n`
	
	`blockPublicAccess.blockPublicPolicy`
		A Boolean value that specifies whether Amazon S3 blocks public bucket policies for the affected bucket. This is a bucket-level, block public access setting for the bucket.
		
		This field isn't available as a filter option on the console.

—	`resourcesAffected.s3Bucket.publicAccess.\n`
	
	`permissionConfiguration.bucketLevelPermissions.\n`
	
	`blockPublicAccess.ignorePublicAcls`
		A Boolean value that specifies whether Amazon S3 ignores public ACLs for the affected bucket and objects in the bucket. This is a bucket-level, block public access setting for the bucket.
		
		This field isn't available as a filter option on the console.

—	`resourcesAffected.s3Bucket.publicAccess.\n`
	
	`permissionConfiguration.bucketLevelPermissions.\n`
	
	`blockPublicAccess.restrictPublicBuckets`
		A Boolean value that specifies whether Amazon S3 restricts public bucket policies for the affected bucket. This is a bucket-level, block public access setting for the bucket.
		
		This field isn't available as a filter option on the console.

Field	JSON field	Description

—	`resourcesAffected.s3Bucket.publicAccess.\n`
	
	`permissionConfiguration.bucketLevelPermissions.\n`
	
	`bucketPolicy.allowsPublicReadAccess`
		A Boolean value that specifies whether the affected bucket's policy allows the general public to have read access to the bucket.
		
		This field isn't available as a filter option on the console.

—	`resourcesAffected.s3Bucket.publicAccess.\n`
	
	`permissionConfiguration.bucketLevelPermissions.\n`
	
	`bucketPolicy.allowsPublicWriteAccess`
		A Boolean value that specifies whether the affected bucket's policy allows the general public to have write access to the bucket.
		
		This field isn't available as a filter option on the console.

S3 bucket tag key*	`resourcesAffected.s3Bucket.tags.key`	A tag key that's associated with the affected bucket.
S3 bucket tag value*	`resourcesAffected.s3Bucket.tags.value`	A tag value that's associated with the affected bucket.

* To specify multiple values for this field on the console, add a condition that uses the field and specifies a distinct value for the filter, and then repeat that step for each additional value. To do this with the API, use an array that lists the values to use for the filter.

S3 object

This table lists and describes fields that you can use to filter findings based on characteristics of the S3 object that a finding applies to.

In the table, the **Field** column indicates the name of the field on the Amazon Macie console. The **JSON field** column uses dot notation to indicate the name of the field in JSON representations of findings and the Amazon Macie API. (Longer JSON field names use the newline character sequence (\n) to improve readability.) The **Description** column provides a brief description of the data that the field stores, and indicates any requirements for filter values. The table is sorted in ascending alphabetical order by field, and then by JSON field.

Field	JSON field	Description

S3 object encryption KMS key id*	`resourcesAffected.s3Object.\n`
	
	`serverSideEncryption.kmsMasterKeyId`
		The Amazon Resource Name (ARN) or unique identifier (key ID) for the AWS KMS key that was used to encrypt the affected object.

S3 object encryption type	`resourcesAffected.s3Object.\n`
	
	`serverSideEncryption.encryptionType`
		The server-side encryption algorithm that was used to encrypt the affected object.
		
		The console provides a list of values to choose from when you add this field to a filter. For a list of valid values for the API, see EncryptionType in the *Amazon Macie API Reference*.

—	`resourcesAffected.s3Object.extension`	The file name extension of the affected object. For objects that don't have a file

Field	JSON field	Description
		name extension, specify "" as the value for the filter. This field isn't available as a filter option on the console.
—	`resourcesAffected.s3Object.lastModified`	The date and time when the affected object was created or last changed, whichever is latest. This field isn't available as a filter option on the console. With the API, you can use this field to define a time range for a filter.
S3 object key*	`resourcesAffected.s3Object.key`	The full name (*key*) of the affected object, including the object's prefix if applicable.
—	`resourcesAffected.s3Object.path`	The full path to the affected object, including the name of the affected bucket and the object's name (*key*). This field isn't available as a filter option on the console.

Field	JSON field	Description
S3 object public access	`resourcesAffected.s3Object.publicAccess`	A Boolean value that specifies whether the affected object is publicly accessible based on a combination of permission settings that apply to the object. The console provides a list of values to choose from when you add this field to a filter.
S3 object tag key*	`resourcesAffected.s3Object.tags.key`	A tag key that's associated with the affected object.
S3 object tag value*	`resourcesAffected.s3Object.tags.value`	A tag value that's associated with the affected object.

* To specify multiple values for this field on the console, add a condition that uses the field and specifies a distinct value for the filter, and then repeat that step for each additional value. To do this with the API, use an array that lists the values to use for the filter.

Fields for policy findings

The following table lists and describes fields that you can use to filter policy findings. These fields store data that's specific to policy findings.

In the table, the **Field** column indicates the name of the field on the Amazon Macie console. The **JSON field** column uses dot notation to indicate the name of the field in JSON representations of findings and the Amazon Macie API. (Longer JSON field names use the newline character sequence (\n) to improve readability.) The **Description** column provides a brief description of the data that the field stores, and indicates any requirements for filter values. The table is sorted in ascending alphabetical order by field, and then by JSON field.

Field	JSON field	Description
Action type	`policyDetails.action.actionType`	The type of action that produced the finding. The only valid value for this field is `AWS_API_CALL`.
API call name*	`policyDetails.action.apiCallDetails.api`	The name of the operation that was invoked most recently and produced the finding.
API service name*	`policyDetails.action.apiCallDetails.apiServiceName`	The URL of the AWS service that provides the operation that was invoked and produced the finding—for example, `s3.amazonaws.com`.
—	`policyDetails.action.apiCallDetails.firstSeen`	The first date and time when any operation was invoked and produced the finding. This field isn't available as a filter option on the console. With the API, you can use this field to define a time range for a filter.
—	`policyDetails.action.apiCallDetails.lastSeen`	The most recent date and time when the specified operation (**API call name** or `api`) was invoked and produced the finding. This field isn't available as a filter option on the console. With the API, you can use this

Field	JSON field	Description
		field to define a time range for a filter.
—	`policyDetails.actor.domainDetails.domainName`	The domain name of the device that was used to perform the action. This field isn't available as a filter option on the console.
IP city*	`policyDetails.actor.ipAddressDetails.ipCity.name`	The name of the originating city for the IP address of the device that was used to perform the action.
IP country*	`policyDetails.actor.ipAddressDetails.ipCountry.name`	The name of the originating country for the IP address of the device that was used to perform the action—for example, `United States`.
—	`policyDetails.actor.ipAddressDetails.ipOwner.asn`	The Autonomous System Number (ASN) for the autonomous system that included the IP address of the device that was used to perform the action. This field isn't available as a filter option on the console.
IP owner ASN org*	`policyDetails.actor.ipAddressDetails.ipOwner.asnOrg`	The organization identifier that's associated with the ASN for the autonomous system that included the IP address of the device that was used to perform the action.

Field	JSON field	Description
IP owner ISP*	`policyDetails.actor.ipAddressDetails.ipOwner.isp`	The name of the internet service provider (ISP) that owned the IP address of the device that was used to perform the action.
IP V4 address*	`policyDetails.actor.ipAddressDetails.ipAddressV4`	The Internet Protocol version 4 (IPv4) address of the device that was used to perform the action.

—	`policyDetails.actor.userIdentity.\n`
	
	`assumedRole.accessKeyId`
		For an action performed with temporary security credentials that were obtained using the `AssumeRole` operation of the AWS STS API, the AWS access key ID that identifies the credentials.
		
		This field isn't available as a filter option on the console.

User identity assumed role account id*	`policyDetails.actor.userIdentity.\n`
	
	`assumedRole.accountId`
		For an action performed with temporary security credentials that were obtained using the `AssumeRole` operation of the AWS STS API, the unique identifier for the AWS account that owns the entity that was used to get the credentials.

Field	JSON field	Description

User identity assumed role principal id*	`policyDetails.actor.userIdentity.\n`
	
	`assumedRole.principalId`
		For an action performed with temporary security credentials that were obtained using the `AssumeRole` operation of the AWS STS API, the unique identifier for the entity that was used to get the credentials.

User identity assumed role session ARN*	`policyDetails.actor.userIdentity.\n`
	
	`assumedRole.arn`
		For an action performed with temporary security credentials that were obtained using the `AssumeRole` operation of the AWS STS API, the Amazon Resource Name (ARN) of the source account, IAM user, or role that was used to get the credentials.

—	`policyDetails.actor.userIdentity.\n`
	
	`assumedRole.sessionContext.sessionIssuer.type`
		For an action performed with temporary security credentials that were obtained using the `AssumeRole` operation of the AWS STS API, the source of the temporary security credentials— for example, `Root`, `IAMUser`, or `Role`.
		
		This field isn't available as a filter option on the console.

Field	JSON field	Description

—	`policyDetails.actor.userIdentity.\n`
	
	`assumedRole.sessionContext.sessionIssuer.userName`
		For an action performed with temporary security credentials that were obtained using the `AssumeRole` operation of the AWS STS API, the name or alias of the user or role that issued the session. Note that this value is null if the credentials were obtained from a root account that doesn't have an alias.
		
		This field isn't available as a filter option on the console.

User identity AWS account account id*	`policyDetails.actor.userIdentity.\n`
	
	`awsAccount.accountId`
		For an action performed using the credentials for another AWS account, the unique identifier for the account.

User identity AWS account principal id*	`policyDetails.actor.userIdentity.\n`
	
	`awsAccount.principalId`
		For an action performed using the credentials for another AWS account, the unique identifier for the entity that performed the action.

User identity AWS service invoked by	`policyDetails.actor.userIdentity.\n`
	
	`awsService.invokedBy`
		For an action performed by an account that belongs to an AWS service, the name of the service.

Field	JSON field	Description

—	`policyDetails.actor.userIdentity.\n`
	
	`federatedUser.accessKeyId`
		For an action performed with temporary security credentials that were obtained using the `GetFederationToken` operation of the AWS STS API, the AWS access key ID that identifies the credentials.
		
		This field isn't available as a filter option on the console.

User identity federated session ARN*	`policyDetails.actor.userIdentity.\n`
	
	`federatedUser.arn`
		For an action performed with temporary security credentials that were obtained using the `GetFederationToken` operation of the AWS STS API, the ARN of the entity that was used to get the credentials.

User identity federated user account id*	`policyDetails.actor.userIdentity.\n`
	
	`federatedUser.accountId`
		For an action performed with temporary security credentials that were obtained using the `GetFederationToken` operation of the AWS STS API, the unique identifier for the AWS account that owns the entity that was used to get the credentials.

Field	JSON field	Description

User identity federated user principal id*	`policyDetails.actor.userIdentity.\n`
	
	`federatedUser.principalId`
		For an action performed with temporary security credentials that were obtained using the `GetFederationToken` operation of the AWS STS API, the unique identifier for the entity that was used to get the credentials.

—	`policyDetails.actor.userIdentity.\n`
	
	`federatedUser.sessionContext.sessionIssuer.type`
		For an action performed with temporary security credentials that were obtained using the `GetFederationToken` operation of the AWS STS API, the source of the temporary security credentials—for example, `Root`, `IAMUser`, or `Role`.
		
		This field isn't available as a filter option on the console.

Field	JSON field	Description

—	`policyDetails.actor.userIdentity.\n`
	
	`federatedUser.sessionContext.sessionIssuer.userName`
		For an action performed with temporary security credentials that were obtained using the `GetFederationToken` operation of the AWS STS API, the name or alias of the user or role that issued the session. Note that this value is null if the credentials were obtained from a root account that doesn't have an alias.
		
		This field isn't available as a filter option on the console.

User identity IAM account id*	`policyDetails.actor.userIdentity.\n`
	
	`iamUser.accountId`
		For an action performed using an IAM user's credentials, the unique identifier for the AWS account that's associated with the IAM user who performed the action.

User identity IAM principal id*	`policyDetails.actor.userIdentity.\n`
	
	`iamUser.principalId`
		For an action performed using an IAM user's credentials, the unique identifier for the IAM user who performed the action.

User identity IAM user name*	`policyDetails.actor.userIdentity.\n`
	
	`iamUser.userName`
		For an action performed using an IAM user's credentials, the username of the IAM user who performed the action.

Field	JSON field	Description
User identity root account id*	`policyDetails.actor.userIdentity.\n root.accountId`	For an action performed using the credentials for your AWS account, the unique ide ntifier for the account.
User identity root principal id*	`policyDetails.actor.userIdentity.\n root.principalId`	For an action performed using the credentials for your AWS account, the unique identifier for the entity that performed the action.
User identity type	`policyDetails.actor.userIdentity.type`	The type of entity that performed the action that produced the finding. The console provides a list of values to choose from when you add this field to a filter. For a list of valid values for the API, see UserIdentityType in the *Amazon Macie API Reference*.

* To specify multiple values for this field on the console, add a condition that uses the field and specifies a distinct value for the filter, and then repeat that step for each additional value. To do this with the API, use an array that lists the values to use for the filter.

Fields for sensitive data findings

The following table lists and describes fields that you can use to filter sensitive data findings. These fields store data that's specific to sensitive data findings.

In the table, the **Field** column indicates the name of the field on the Amazon Macie console. The **JSON field** column uses dot notation to indicate the name of the field in JSON representations of findings and the Amazon Macie API. (Longer JSON field names use the newline character sequence (\n) to improve readability.) The **Description** column provides a brief description of the data that

the field stores, and indicates any requirements for filter values. The table is sorted in ascending alphabetical order by field, and then by JSON field.

Field	JSON field	Description
Custom data identifier ID*	`classificationDetails.result.\n customDataIdentifiers.detections.arn`	The unique identifier for the custom data identifier that detected the data and produced the finding.
Custom data identifier name*	`classificationDetails.result.\n customDataIdentifiers.detections.name`	The name of the custom data identifier that detected the data and produced the finding.
Custom data identifier total count	`classificationDetails.result.\n customDataIdentifiers.detections.count`	The total number of occurrences of data that was detected by custom data identifiers and produced the finding. You can use this field to define a numeric range for a filter.
Job ID*	`classificationDetails.jobId`	The unique identifier for the sensitive data discovery job that produced the finding.
Origin type	`classificationDetails.originType`	How Macie found the sensitive data that produced the finding: `AUTOMATED_SENSITIVE_DATA_DISCOVERY` or `SENSITIVE_DATA_DISCOVERY_JOB`.

Field	JSON field	Description
—	`classificationDetails.result.mimeType`	The type of content, as a MIME type, that the finding applies to—for example, `text/csv` for a CSV file or `application/pdf` for an Adobe Portable Document Format file. This field isn't available as a filter option on the console.
—	`classificationDetails.result.sizeClassified`	The total storage size, in bytes, of the S3 object that the finding applies to. This field isn't available as a filter option on the console. With the API, you can use this field to define a numeric range for a filter.

Field	JSON field	Description
Result status code*	`classificationDetails.result.status.code`	The status of the finding. Valid values are: • `COMPLETE` – Macie completed its analysis of the object. • `PARTIAL` – Macie analyzed only a subset of the data in the object. For example, the object is an archive file that contains files in an unsupported format. • `SKIPPED` – Macie wasn't able to analyze the object. For example, the object is a malformed file.

Sensitive data category	`classificationDetails.result.\n`
	
	`sensitiveData.category`
		The category of sensitive data that was detected and produced the finding.
		
		The console provides a list of values to choose from when you add this field to a filter. In the API, valid values are: `CREDENTIALS`, `FINANCIAL_INFORMATION`, and `PERSONAL_INFORMATION`.

Field	JSON field	Description

Sensitive data detection type	`classificationDetails.result.\n`
	
	`sensitiveData.detections.type`
		The type of sensitive data that was detected and produced the finding. This is the unique identifier for the managed data identifier that detected the data.
		
		The console provides a list of values to choose from when you add this field to a filter. For a list of valid values for both the console and the API, see Quick reference: Managed data identifiers by type.

Sensitive data total count	`classificationDetails.result.\n`
	
	`sensitiveData.detections.count`
		The total number of occurrences of the type of sensitive data that was detected and produced the finding.
		
		You can use this field to define a numeric range for a filter.

* To specify multiple values for this field on the console, add a condition that uses the field and specifies a distinct value for the filter, and then repeat that step for each additional value. To do this with the API, use an array that lists the values to use for the filter.

Creating and applying filters to Macie findings

To identify and focus on findings that have specific characteristics, you can filter findings on the Amazon Macie console and in queries that you submit programmatically using the Amazon Macie API. When you create a filter, you use specific attributes of findings to define criteria for including

or excluding findings from a view or from query results. A *finding attribute* is a field that stores specific data for a finding, such as severity, type, or the name of the resource that a finding applies to.

In Macie, a filter consists of one or more conditions. Each condition, also referred to as a *criterion*, consists of three parts:

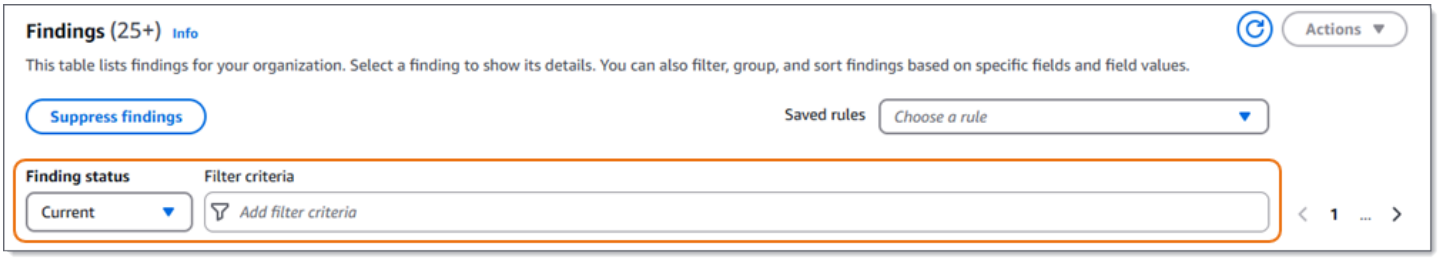
- An attribute-based field, such as **Severity** or **Finding type**.
- An operator, such as *equals* or *not equals*.
- One or more values. The type and number of values depends on the field and operator that you choose.

How you define and apply filter conditions depends on whether you use the Amazon Macie console or the Amazon Macie API.

Topics

Filtering findings by using the Amazon Macie console

If you use the Amazon Macie console to filter findings, Macie provides options to help you choose fields, operators, and values for individual conditions. You access these options by using filter settings on **Findings** pages, as shown in the following image.



By using the **Finding status** menu, you can specify whether to include findings that were suppressed (automatically archived) by a suppression rule. By using the **Filter criteria** box, you can enter filter conditions.

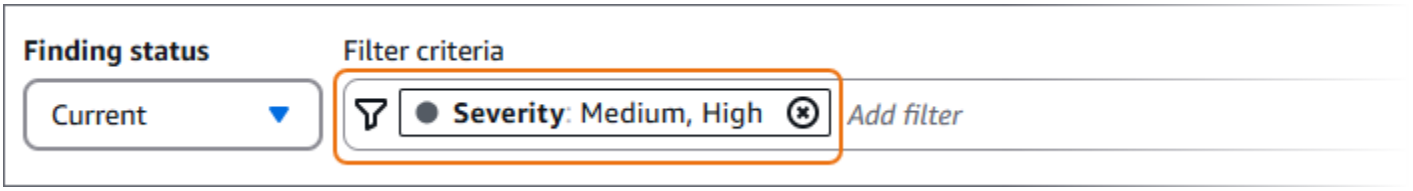
When you place your cursor in the **Filter criteria** box, Macie displays a list of fields that you can use in filter conditions. The fields are organized by logical category. For example, the **Common fields**

category includes fields that apply to any type of finding, and the **Classification fields** category includes fields that apply only to sensitive data findings. The fields are sorted alphabetically within each category.

To add a condition, start by choosing a field from the list. To find a field, browse the complete list, or enter part of the field's name to narrow the list of fields.

Depending on the field that you choose, Macie displays different options. The options reflect the type and nature of the field that you choose. For example, if you choose the **Severity** field, Macie displays a list of values to choose from—**Low**, **Medium**, and **High**. If you choose the **S3 bucket name** field, Macie displays a text box in which you can enter a bucket name. Whichever field you choose, Macie guides you through the steps to add a condition that includes the required settings for the field.

After you add a condition, Macie applies the criteria for the condition and adds the condition to a filter token in the **Filter criteria** box, as shown in the following image.



In this example, the condition is configured to include all medium-severity and high-severity findings, and to exclude all low-severity findings. It returns findings where the value for the **Severity** field *equals* **Medium** or **High**.

> **Tip**
>
> For many fields, you can change a condition's operator from *equals* to *not equals* by choosing the equals icon (●) in the filter token for the condition. If you do this, Macie changes the operator to *not equals* and displays the not equals icon (⃠) in the token. To switch to the *equals* operator again, choose the not equals icon.

As you add more conditions, Macie applies their criteria and adds them to tokens in the **Filter criteria** box. You can refer to the box at any time to determine which

criteria you've applied. To remove a condition, choose the remove condition icon
( )
in the token for the condition.

To filter findings by using the console

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Findings**.
3. (Optional) To first pivot on and review findings by a predefined logical group, choose **By bucket**, **By type**, or **By job** in the navigation pane (under **Findings**). Then choose an item in the table. In the details panel, choose the link for the field to pivot on.
4. (Optional) To display findings that were suppressed by a suppression rule, change the **Filter status** setting. Choose **Archived** to display only suppressed findings, or choose **All** to display both suppressed and unsuppressed findings. To hide suppressed findings, choose **Current**.
5. To add a filter condition:

   a. Place your cursor in the **Filter criteria** box, and then choose the field to use for the condition. For information about the fields that you can use, see Fields for filtering Macie findings.

   b. Enter the appropriate type of value for the field. For detailed information about the different types of values, see Specifying values for fields.

      Array of text (strings)

      For this type of value, Macie often provides a list of values to choose from. If this is the case, select each value that you want to use in the condition.

      If Macie doesn't provide a list of values, enter a complete, valid value for the field. To specify additional values for the field, choose **Apply**, and then add another condition for each additional value.

      Note that values are case sensitive. In addition, you can't use partial values or wildcard characters in values. For example, to filter findings for an S3 bucket named *my-S3-bucket*, enter **`my-S3-bucket`** as the value for the **S3 bucket name** field. If you enter any other value, such as **`my-s3-bucket`** or **`my-S3`**, Macie won't return findings for the bucket.

Boolean

For this type of value, Macie provides a list of values to choose from. Select the value that you want to use in the condition.

Date/Time (time ranges)

For this type of value, use the **From** and **To** boxes to define an inclusive time range:

- To define a fixed time range, use the **From** and **To** boxes to specify the first date and time and the last date and time in the range, respectively.
- To define a relative time range that starts at a certain date and time and ends at the current time, enter the start date and time in the **From** boxes, and delete any text in **To** boxes.
- To define a relative time range that ends at a certain date and time, enter the end date and time in the **To** boxes, and delete any text in the **From** boxes.

Note that time values use 24-hour notation. If you use the date picker to choose dates, you can refine the values by entering text directly in the **From** and **To** boxes.

Number (numeric ranges)

For this type of value, use the **From** and **To** boxes to enter one or more integers that define an inclusive, fixed or relative numeric range.

Text (string) values

For this type of value, enter a complete, valid value for the field.

Note that values are case sensitive. In addition, you can't use partial values or wildcard characters in values. For example, to filter findings for an S3 bucket named *my-S3-bucket*, enter **`my-S3-bucket`** as the value for the **S3 bucket name** field. If you enter any other value, such as **`my-s3-bucket`** or **`my-S3`**, Macie won't return findings for the bucket.

c. When you finish adding values for the field, choose **Apply**. Macie applies the filter criteria and adds the condition to a filter token in the **Filter criteria** box.

6. Repeat step 5 for each additional condition that you want to add.
7. To remove a condition, choose the remove condition icon (⊗) in the filter token for the condition.

8. To change a condition, remove the condition by choosing the remove condition icon () in the filter token for the condition. Then repeat step 5 to add a condition with the correct settings.

> **Tip**
>
> If you want to subsequently use this set of conditions again, you can save the set as a filter rule. To do this, choose **Save rule** in the **Filter criteria** box. Then enter a name and, optionally, a description for the rule. When you finish, choose **Save**.

Filtering findings programmatically with the Amazon Macie API

To filter findings programmatically, specify filter criteria in queries that you submit using the ListFindings or GetFindingStatistics operation of the Amazon Macie API. The **ListFindings** operation returns an array of finding IDs, one ID for each finding that matches the filter criteria. The **GetFindingStatistics** operation returns aggregated statistical data about all the findings that match the filter criteria, grouped by a field that you specify in your request.

Note that the **ListFindings** and **GetFindingStatistics** operations are different from operations that you use to suppress findings. Unlike suppression operations, which also specify filter criteria, the **ListFindings** and **GetFindingStatistics** operations only query findings data. They don't perform any action on findings that match filter criteria. To suppress findings, use the CreateFindingsFilter operation of the Amazon Macie API.

To specify filter criteria in a query, include a map of filter conditions in your request. For each condition, specify a field, an operator, and one or more values for the field. The type and number of values depends on the field and operator that you choose. For information about the fields, operators, and types of values that you can use in a condition, see Fields for filtering Macie findings, Using operators in conditions, and Specifying values for fields.

The following examples show you how to specify filter criteria in queries that you submit using the AWS Command Line Interface (AWS CLI). You can also do this by using a current version of another AWS command line tool or an AWS SDK, or by sending HTTPS requests directly to Macie. For information about AWS tools and SDKs, see Tools to Build on AWS.

Examples

- Example 1: Filter findings based on severity
- Example 2: Filter findings based on sensitive data category
- Example 3: Filter findings based on a fixed time range
- Example 4: Filter findings based on suppression status
- Example 5: Filter findings based on multiple fields and types of values

The examples use the list-findings command. If an example runs successfully, Macie returns a `findingIds` array. The array lists the unique identifier for each finding that matches the filter criteria, as shown in the following example.

```
{
    "findingIds": [
        "1f1c2d74db5d8caa76859ec52example",
        "6cfa9ac820dd6d55cad30d851example",
        "702a6fd8750e567d1a3a63138example",
        "826e94e2a820312f9f964cf60example",
        "274511c3fdcd87010a19a3a42example"
    ]
}
```

If no findings match the filter criteria, Macie returns an empty `findingIds` array.

```
{
    "findingIds": []
}
```

Example 1: Filter findings based on severity

This example retrieves finding IDs for all high-severity and medium-severity findings in the current AWS Region.

For Linux, macOS, or Unix:

```
$ aws macie2 list-findings --finding-criteria '{"criterion":{"severity.description":
{"eq":["High","Medium"]}}}'
```

For Microsoft Windows:

```
C:\> aws macie2 list-findings --finding-criteria={\"criterion\":
{\"severity.description\":{\"eq\":[\"High\",\"Medium\"]}}}
```

Where:

- *severity.description* specifies the JSON name of the **Severity** field.
- *eq* specifies the *equals* operator.
- *High* and *Medium* are an array of enumerated values for the **Severity** field.

Example 2: Filter findings based on sensitive data category

This example retrieves finding IDs for all sensitive data findings that are in the current Region and report occurrences of financial information (and no other categories of sensitive data) in S3 objects.

For Linux, macOS, or Unix, using the backslash (\) line-continuation character to improve readability:

```
$ aws macie2 list-findings \
--finding-criteria '{"criterion":
{"classificationDetails.result.sensitiveData.category":{"eqExactMatch":
["FINANCIAL_INFORMATION"]}}}'
```

For Microsoft Windows, using the caret (^) line-continuation character to improve readability:

```
C:\> aws macie2 list-findings ^
--finding-criteria={\"criterion\":
{\"classificationDetails.result.sensitiveData.category\":{\"eqExactMatch\":
[\"FINANCIAL_INFORMATION\"]}}}
```

Where:

- *classificationDetails.result.sensitiveData.category* specifies the JSON name of the **Sensitive data category** field.
- *eqExactMatch* specifies the *equals exact match* operator.
- *FINANCIAL_INFORMATION* is an enumerated value for the **Sensitive data category** field.

Example 3: Filter findings based on a fixed time range

This example retrieves finding IDs for all findings that are in the current Region and were created between 07:00 UTC October 5, 2020, and 07:00 UTC November 5, 2020 (inclusively).

For Linux, macOS, or Unix:

```
$ aws macie2 list-findings --finding-criteria '{"criterion":{"createdAt":
{"gte":1601881200000,"lte":1604559600000}}}'
```

For Microsoft Windows:

```
C:\> aws macie2 list-findings --finding-criteria={\"criterion\":{\"createdAt\":
{\"gte\":1601881200000,\"lte\":1604559600000}}}
```

Where:

- *createdAt* specifies the JSON name of the **Created at** field.
- *gte* specifies the *greater than or equal to* operator.
- *1601881200000* is the first date and time (as a Unix timestamp in milliseconds) in the time range.
- *lte* specifies the *less than or equal to* operator.
- *1604559600000* is the last date and time (as a Unix timestamp in milliseconds) in the time range.

Example 4: Filter findings based on suppression status

This example retrieves finding IDs for all findings that are in the current Region and were suppressed (automatically archived) by a suppression rule.

For Linux, macOS, or Unix:

```
$ aws macie2 list-findings --finding-criteria '{"criterion":{"archived":{"eq":
["true"]}}}'
```

For Microsoft Windows:

```
C:\> aws macie2 list-findings --finding-criteria={\"criterion\":{\"archived\":{\"eq\":
[\"true\"]}}}
```

Where:

- *archived* specifies the JSON name of the **Archived** field.
- *eq* specifies the *equals* operator.
- *true* is a Boolean value for the **Archived** field.

Example 5: Filter findings based on multiple fields and types of values

This example retrieves finding IDs for all sensitive data findings that are in the current Region and match the following criteria: were created between 07:00 UTC October 5, 2020, and 07:00 UTC November 5, 2020 (exclusively); report occurrences of financial data and no other categories of sensitive data in S3 objects; and weren't suppressed (automatically archived) by a suppression rule.

For Linux, macOS, or Unix, using the backslash (\) line-continuation character to improve readability:

```
$ aws macie2 list-findings \
--finding-criteria '{"criterion":{"createdAt":
{"gt":1601881200000,"lt":1604559600000},"classificationDetails.result.sensitiveData.category":
{"eqExactMatch":["FINANCIAL_INFORMATION"]},"archived":{"eq":["false"]}}}'
```

For Microsoft Windows, using the caret (^) line-continuation character to improve readability:

```
C:\> aws macie2 list-findings ^
--finding-criteria={\"criterion\":{\"createdAt\":{\"gt\":1601881200000,
\"lt\":1604559600000},\"classificationDetails.result.sensitiveData.category\":
{\"eqExactMatch\":[\"FINANCIAL_INFORMATION\"]},\"archived\":{\"eq\":[\"false\"]}}}
```

Where:

- *createdAt* specifies the JSON name of the **Created at** field, and:
  - *gt* specifies the *greater than or equal to* operator.
  - *1601881200000* is the first date and time (as a Unix timestamp in milliseconds) in the time range.
  - *lt* specifies the *less than or equal to* operator.
  - *1604559600000* is the last date and time (as a Unix timestamp in milliseconds) in the time range.

- `classificationDetails.result.sensitiveData.category` specifies the JSON name of the **Sensitive data category** field, and:
  - `eqExactMatch` specifies the *equals exact match* operator.
  - `FINANCIAL_INFORMATION` is an enumerated value for the field.
- `archived` specifies the JSON name of the **Archived** field, and:
  - `eq` specifies the *equals* operator.
  - `false` is a Boolean value for the field.

Defining filter rules for Macie findings

To perform consistent analysis of findings, you can create and apply filter rules. A *filter rule* is a set of filter criteria that you create and save to use again when you review findings on the Amazon Macie console. Filter rules can help you perform repeated, consistent analysis of findings that have specific characteristics. For example, you might create one filter rule for analyzing all high-severity sensitive data findings that report specific types of sensitive data. You might create another filter rule for analyzing all high-severity policy findings for Amazon Simple Storage Service (Amazon S3) buckets that store unencrypted objects.

When you create a filter rule, you use specific attributes of findings to define criteria for including or excluding findings from a view. A *finding attribute* is a field that stores specific data for a finding, such as severity, type, or the name of the S3 bucket that a finding applies to. You also specify a name, and, optionally, a description of the rule. To then analyze findings that match the criteria of the rule, choose the rule. Macie applies the rule's criteria and displays only those findings that match the criteria. Macie also displays the criteria to help you determine which criteria it applied.

Note that filter rules are different from suppression rules. A *suppression rule* is a set of filter criteria that you create and save to automatically archive findings that match the criteria of the rule. Although both types of rules store and apply filter criteria, a filter rule doesn't perform any action on findings that match the rule's criteria. Instead, a filter rule only determines which findings appear on the console after you apply the rule. For information about suppression rules, see Suppressing findings.

Topics

- Creating a filter rule for Macie findings
- Applying a filter rule to Macie findings
- Changing a filter rule for Macie findings

- Deleting a filter rule for Macie findings

Creating a filter rule for Macie findings

A *filter rule* is a set of filter criteria that you create and save to use again when you review findings on the Amazon Macie console. Filter rules can help you perform repeated, consistent analysis of findings that have specific characteristics. For example, you might create a filter rule for analyzing all high-severity sensitive data findings that report occurrences of sensitive data in particular Amazon Simple Storage Service (Amazon S3) buckets. You can then apply that filter rule each time you want to identify and analyze findings that have the specified characteristics.

When you create a filter rule, you specify filter criteria, a name, and, optionally, a description of the rule. For the filter criteria, you use specific attributes of findings to specify whether to include or exclude findings from a view. A *finding attribute* is a field that stores specific data for a finding, such as severity, type, or the name of the resource that a finding applies to. Filter criteria consist of one or more conditions. Each condition, also referred to as a *criterion*, consists of three parts:

- An attribute-based field, such as **Severity** or **Finding type**.
- An operator, such as *equals* or *not equals*.
- One or more values. The type and number of values depends on the field and operator that you choose.

After you create and save a filter rule, you apply its filter criteria by choosing the rule. Macie then uses the criteria to determine which findings to display. Macie also displays the criteria to help you determine which criteria you applied.

Note that filter rules are different from suppression rules. A *suppression rule* is a set of filter criteria that you create and save to automatically archive findings that match the criteria of the rule. Although both types of rules store and apply filter criteria, a filter rule doesn't perform any action on findings that match the rule's criteria. Instead, a filter rule only determines which findings appear on the console after you apply the rule. For information about suppression rules, see Suppressing findings.

To create a filter rule for findings

You can create a filter rule by using the Amazon Macie console or the Amazon Macie API.

Console

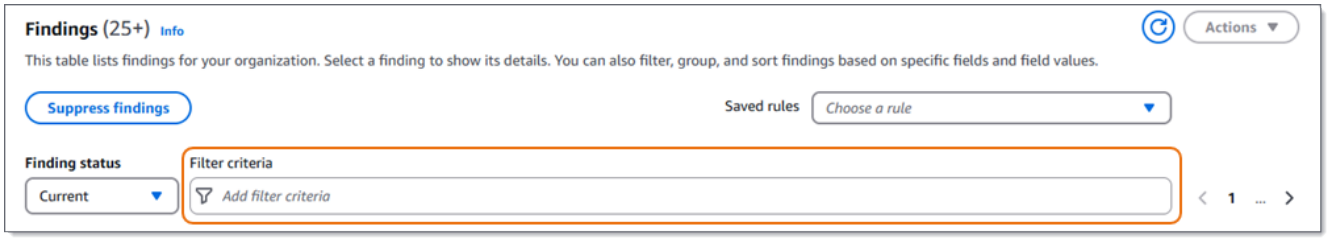
Follow these steps to create a filter rule by using the Amazon Macie console.

To create a filter rule

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Findings**.

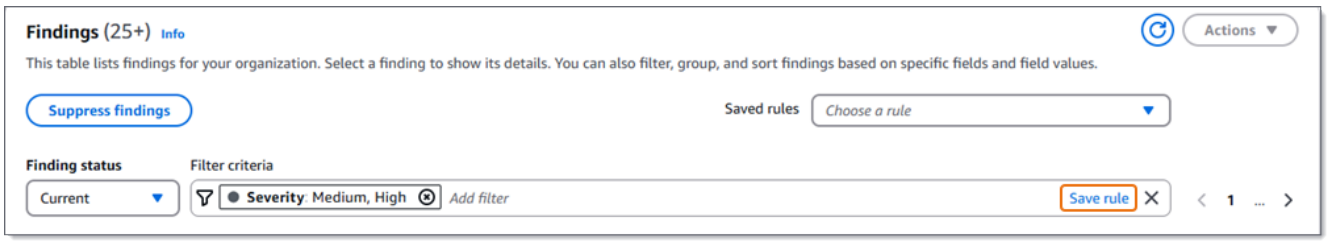
> **Tip**
>
> To use an existing filter rule as a starting point, choose the rule from the **Saved rules** list.
>
> You can also streamline creation of a rule by first pivoting and drilling down on findings by a predefined logical group. If you do this, Macie automatically creates and applies the appropriate filter conditions, which can be a helpful starting point for creating a rule. To do this, choose **By bucket**, **By type**, or **By job** in the navigation pane (under **Findings**). Then choose an item in the table. In the details panel, choose the link for the field to pivot on.

3. In the **Filter criteria** box, add conditions that define the filter criteria for the rule.



   To learn how to add filter conditions, see Creating and applying filters to Macie findings.

4. When you finish defining filter criteria for the rule, choose **Save rule** in the **Filter criteria** box.



5. Under **Filter rule**, enter a name and, optionally, a description of the rule.
6. Choose **Save**.

API

To create a filter rule programmatically, use the CreateFindingsFilter operation of the Amazon Macie API and specify the appropriate values for the required parameters:

- For the `action` parameter, specify `NOOP` to ensure that Macie doesn't suppress (automatically archive) findings that match the criteria of the rule.
- For the `criterion` parameter, specify a map of conditions that define the filter criteria for the rule.

  In the map, each condition should specify a field, an operator, and one or more values for the field. The type and number of values depends on the field and operator that you choose. For information about the fields, operators, and types of values that you can use in a condition, see: Fields for filtering Macie findings, Using operators in conditions, and Specifying values for fields.

To create a filter rule by using the AWS Command Line Interface (AWS CLI), run the create-findings-filter command and specify the appropriate values for the required parameters. The following examples create a filter rule that returns all sensitive data findings that are in the current AWS Region and report occurrences of personal information (and no other categories of sensitive data) in S3 objects.

This example is formatted for Linux, macOS, or Unix, and it uses the backslash (\) line-continuation character to improve readability.

```
$ aws macie2 create-findings-filter \
--action NOOP \
--name my_filter_rule \
--finding-criteria '{"criterion":
{"classificationDetails.result.sensitiveData.category":{"eqExactMatch":
["PERSONAL_INFORMATION"]}}}'
```

This example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws macie2 create-findings-filter ^
--action NOOP ^
--name my_filter_rule ^
```

```
--finding-criteria={\"criterion\":
{\"classificationDetails.result.sensitiveData.category\":{\"eqExactMatch\":
[\"PERSONAL_INFORMATION\"]}}}
```

Where:

- *my_filter_rule* is the custom name for the rule.
- `criterion` is a map of filter conditions for the rule:
  - *classificationDetails.result.sensitiveData.category* is the JSON name of the **Sensitive data category** field.
  - *eqExactMatch* specifies the *equals exact match* operator.
  - *PERSONAL_INFORMATION* is an enumerated value for the **Sensitive data category** field.

If the command runs successfully, you receive output similar to the following.

```
{
    "arn": "arn:aws:macie2:us-west-2:123456789012:findings-filter/9b2b4508-
aa2f-4940-b347-d1451example",
    "id": "9b2b4508-aa2f-4940-b347-d1451example"
}
```

Where `arn` is the Amazon Resource Name (ARN) of the filter rule that was created, and `id` is the unique identifier for the rule.

For additional examples of filter criteria, see Filtering findings programmatically with the Amazon Macie API.

Applying a filter rule to Macie findings

When you apply a filter rule, Amazon Macie uses the rule's criteria to determine which findings to include or exclude from your view of findings on the console. Macie also displays the criteria to help you determine which criteria you applied.

> **Tip**
>
> Although filter rules are designed for use with the Amazon Macie console, you can use a rule's criteria to query findings data programmatically with the Amazon Macie API. To do this, retrieve the filter criteria for the rule, and then add the criteria to your query. To

retrieve the criteria, use the GetFindingsFilter operation. To then identify findings that match the criteria, use the ListFindings operation and specify the criteria in your query. For information about specifying filter criteria in a query, see Creating and applying filters to Macie findings.

To apply a filter rule to findings

Follow these steps to filter findings on the Amazon Macie console by applying a filter rule.

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Findings**.
3. In the **Saved rules** list, choose the filter rule that you want to apply. Macie applies the rule's criteria and displays the criteria in the **Filter criteria** box.
4. To refine the criteria, use the **Filter criteria** box to add or remove filter conditions. If you do this, your changes won't affect the settings for the rule. Macie saves your changes only if you explicitly save them as a new rule.
5. To apply a different filter rule, repeat step 3.

After you apply a filter rule, you can quickly remove all of its filter criteria from your view. To do this, choose the **X** in the **Filter criteria** box.

Changing a filter rule for Macie findings

After you create a filter rule, you can refine its criteria and change other settings for the rule. A *filter rule* is a set of filter criteria that you create and save to use again when you review findings on the Amazon Macie console. Filter rules can help you perform repeated, consistent analysis of findings that have specific characteristics. Each rule consists of a set of filter criteria, a name, and, optionally, a description.

In addition to changing the filter criteria or other settings for a rule, you can assign tags to a rule. A *tag* is a label that you define and assign to certain types of AWS resources. Each tag consists of a required tag key and an optional tag value. Tags can help you identify, categorize, and manage resources in different ways, such as by purpose, owner, environment, or other criteria. To learn more, see Tagging Macie resources.

To change a filter rule for findings

To assign tags or change the settings for a filter rule, you can use the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to assign tags or change the settings for a filter rule by using the Amazon Macie console.

To change a filter rule

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Findings**.
3. In the **Saved rules** list, choose the edit icon () next to the filter rule that you want to change or assign tags to.
4. Do any of the following:
   - To change the filter criteria of the rule, use the **Filter criteria** box. In the box, enter conditions for the criteria that you want. To learn how, see Creating and applying filters to Macie findings.
   - To change the name of the rule, enter a new name in the **Name** box under **Filter rule**.
   - To change the description of the rule, enter a new description in the **Description** box under **Filter rule**.
   - To assign tags to the rule, choose **Manage tags** under **Filter rule**. Then add, review, and change the tags as necessary. A rule can have as many as 50 tags.
5. When you finish making changes, choose **Save**.

API

To change a filter rule programmatically, use the UpdateFindingsFilter operation of the Amazon Macie API. When you submit your request, use the supported parameters to specify a new value for each setting that you want to change.

For the `id` parameter, specify the unique identifier for the rule to change. You can get this identifier by using the ListFindingsFilter operation to retrieve a list of filter and suppression rules for your account. If you're using the AWS Command Line Interface (AWS CLI), run the list-findings-filters command to retrieve this list.

To change a filter rule by using the AWS CLI, run the update-findings-filter command and use the supported parameters to specify a new value for each setting that you want to change. For example, the following command changes the name of an existing filter rule.

```
C:\> aws macie2 update-findings-filter --id 9b2b4508-aa2f-4940-b347-d1451example --
name personal_information_only
```

Where:

- *9b2b4508-aa2f-4940-b347-d1451example* is the unique identifier for the rule.
- *personal_information_only* is the new name for the rule.

If the command runs successfully, you receive output similar to the following.

```
{
    "arn": "arn:aws:macie2:us-west-2:123456789012:findings-filter/9b2b4508-
aa2f-4940-b347-d1451example",
    "id": "9b2b4508-aa2f-4940-b347-d1451example"
}
```

Where `arn` is the Amazon Resource Name (ARN) of the rule that was changed, and `id` is the unique identifier for the rule.

Similarly, the following example converts a suppression rule to a filter rule by changing the value for the `action` parameter from `ARCHIVE` to `NOOP`.

```
C:\> aws macie2 update-findings-filter --id 8a1c3508-aa2f-4940-b347-d1451example --
action NOOP
```

Where:

- *8a1c3508-aa2f-4940-b347-d1451example* is the unique identifier for the rule.
- *NOOP* is the new action for Macie to perform on findings that match the criteria of the rule—perform no action (don't suppress the findings).

If the command runs successfully, you receive output similar to the following:

```
{
```

```
    "arn": "arn:aws:macie2:us-west-2:123456789012:findings-filter/8a1c3508-
aa2f-4940-b347-d1451example",
    "id": "8a1c3508-aa2f-4940-b347-d1451example"
}
```

Where `arn` is the Amazon Resource Name (ARN) of the rule that was changed, and `id` is the unique identifier for the rule.

Deleting a filter rule for Macie findings

If you create a filter rule, you can delete it at any time. A *filter rule* is a set of filter criteria that you create and save to use again when you review findings on the Amazon Macie console. If you delete a filter rule, your change doesn't affect findings that match the rule's criteria. A filter rule only determines which findings appear on the console after you apply the rule.

To delete a filter rule for findings

You can delete a filter rule by using the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to delete a filter rule by using the Amazon Macie console.

To delete a filter rule

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Findings**.
3. In the **Saved rules** list, choose the edit icon () next to the filter rule that you want to delete.
4. Under **Filter rule**, choose **Delete**.

API

To delete a filter rule programmatically, use the DeleteFindingsFilter operation of the Amazon Macie API. For the `id` parameter, specify the unique identifier for the filter rule to delete. You can get this identifier by using the ListFindingsFilter operation to retrieve a list of filter and suppression rules for your account. If you're using the AWS Command Line Interface (AWS CLI), run the list-findings-filters command to retrieve this list.

To delete a filter rule by using the AWS CLI, run the delete-findings-filter command. For example:

```
C:\> aws macie2 delete-findings-filter --id 9b2b4508-aa2f-4940-b347-d1451example
```

Where *9b2b4508-aa2f-4940-b347-d1451example* is the unique identifier for the filter rule to delete.

If the command runs successfully, Macie returns an empty HTTP 200 response. Otherwise, Macie returns an HTTP 4*xx* or 500 response that indicates why the operation failed.

Investigating sensitive data with Macie findings

When you run sensitive data discovery jobs or Amazon Macie performs automated sensitive data discovery, Macie captures details about the location of each occurrence of sensitive data that it finds in Amazon Simple Storage Service (Amazon S3) objects. This includes sensitive data that Macie detects using managed data identifiers, and data that matches the criteria of custom data identifiers that you configure a job or Macie to use.

With sensitive data findings, you can review these details for as many as 15 occurrences of sensitive data that Macie finds in individual S3 objects. The details provide insight into the breadth of the categories and types of sensitive data that specific S3 buckets and objects might contain. They can help you locate individual occurrences of sensitive data in objects, and determine whether to perform a deeper investigation of specific buckets and objects.

For additional insight, you can optionally configure and use Macie to retrieve samples of sensitive data that Macie reports in individual findings. The samples can help you verify the nature of the sensitive data that Macie found. They can also help you tailor your investigation of an affected S3 bucket and object. If you choose to retrieve sensitive data samples for a finding, Macie uses data in the finding to locate 1-10 occurrences of each type of sensitive data reported by the finding. Macie then extracts those occurrences of sensitive data from the affected object and displays the data for you to review.

If an S3 object contains many occurrences of sensitive data, a finding can also help you navigate to the corresponding sensitive data discovery result. Unlike a sensitive data finding, a sensitive data discovery result provides detailed location data for as many as 1,000 occurrences of each type of sensitive data that Macie finds in an object. Macie uses the same schema for location data

in sensitive data findings and sensitive data discovery results. To learn more about sensitive data discovery results, see Storing and retaining sensitive data discovery results.

The topics in this section explain how to locate and optionally retrieve occurrences of sensitive data reported by sensitive data findings. They also explain the schema that Macie uses to report the location of individual occurrences of sensitive data that Macie finds.

Topics

Locating sensitive data with Macie findings

When you run sensitive data discovery jobs or Amazon Macie performs automated sensitive data discovery, Macie performs a deep inspection of the latest version of each Amazon Simple Storage Service (Amazon S3) object that it analyzes. For each job run or analysis cycle, Macie also uses a *depth-first search* algorithm to populate the resulting findings with details about the location of specific occurrences of sensitive data that Macie finds in S3 objects. These occurrences provide insight into the categories and types of sensitive data that an affected S3 bucket and object might contain. The details can help you locate individual occurrences of sensitive data in objects, and determine whether to perform a deeper investigation of specific buckets and objects.

With sensitive data findings, you can determine the location of as many as 15 occurrences of sensitive data that Macie found in an affected S3 object. This includes sensitive data that Macie detected using managed data identifiers, and data that matches the criteria of custom data identifiers that you configured a job or Macie to use.

A sensitive data finding can provide details such as:

- The column and row number for a cell or field in a Microsoft Excel workbook, CSV file, or TSV file.
- The path to a field or array in a JSON or JSON Lines file.
- The line number for a line in a non-binary text file other than a CSV, JSON, JSON Lines, or TSV file—for example, an HTML, TXT, or XML file.
- The page number for a page in an Adobe Portable Document Format (PDF) file.
- The record index and the path to a field in a record in an Apache Avro object container or Apache Parquet file.

You can access these details by using the Amazon Macie console or the Amazon Macie API. You can also access these details in findings that Macie publishes to other AWS services, both Amazon EventBridge and AWS Security Hub. To learn about the JSON structures that Macie uses to report these details, see Schema for reporting the location of sensitive data. To learn how to access the details in findings that Macie publishes to other AWS services, see Monitoring and processing findings.

If an S3 object contains many occurrences of sensitive data, you can also use a finding to navigate to its corresponding sensitive data discovery result. Unlike a sensitive data finding, a sensitive data discovery result provides detailed location data for as many as 1,000 occurrences of each type of sensitive data that Macie found in an object. If an S3 object is an archive file, such as a .tar or .zip file, this includes occurrences of sensitive data in individual files that Macie extracted from the archive. (Macie doesn't include this information in sensitive data findings.) To learn more about sensitive data discovery results, see Storing and retaining sensitive data discovery results. Macie uses the same schema for location data in sensitive data findings and sensitive data discovery results.

To locate sensitive data with findings

To locate occurrences of sensitive data reported by a finding, you can use the Amazon Macie console or the Amazon Macie API. To do this programmatically, use the GetFindings operation. If a finding includes details about the location of one or more occurrences of a specific type of sensitive data, `occurrences` objects in the finding provide these details. For more information, see Schema for reporting the location of sensitive data.

To locate occurrences of sensitive data by using the console, follow these steps.

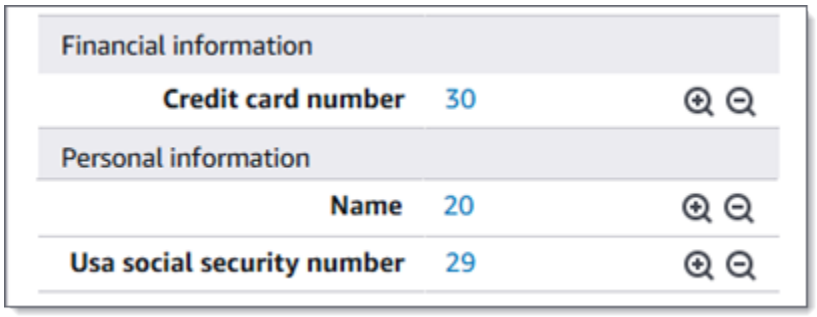
1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Findings**.

   > **Tip**
   >
   > You can quickly display all the findings from a particular sensitive data discovery job. To do this, choose **Jobs** in the navigation pane, and then choose the name of the job. At the top of the details panel, choose **Show results**, and then choose **Show findings**.

3. On the **Findings** page, choose the finding for the sensitive data that you want to locate. The details panel displays information for the finding.

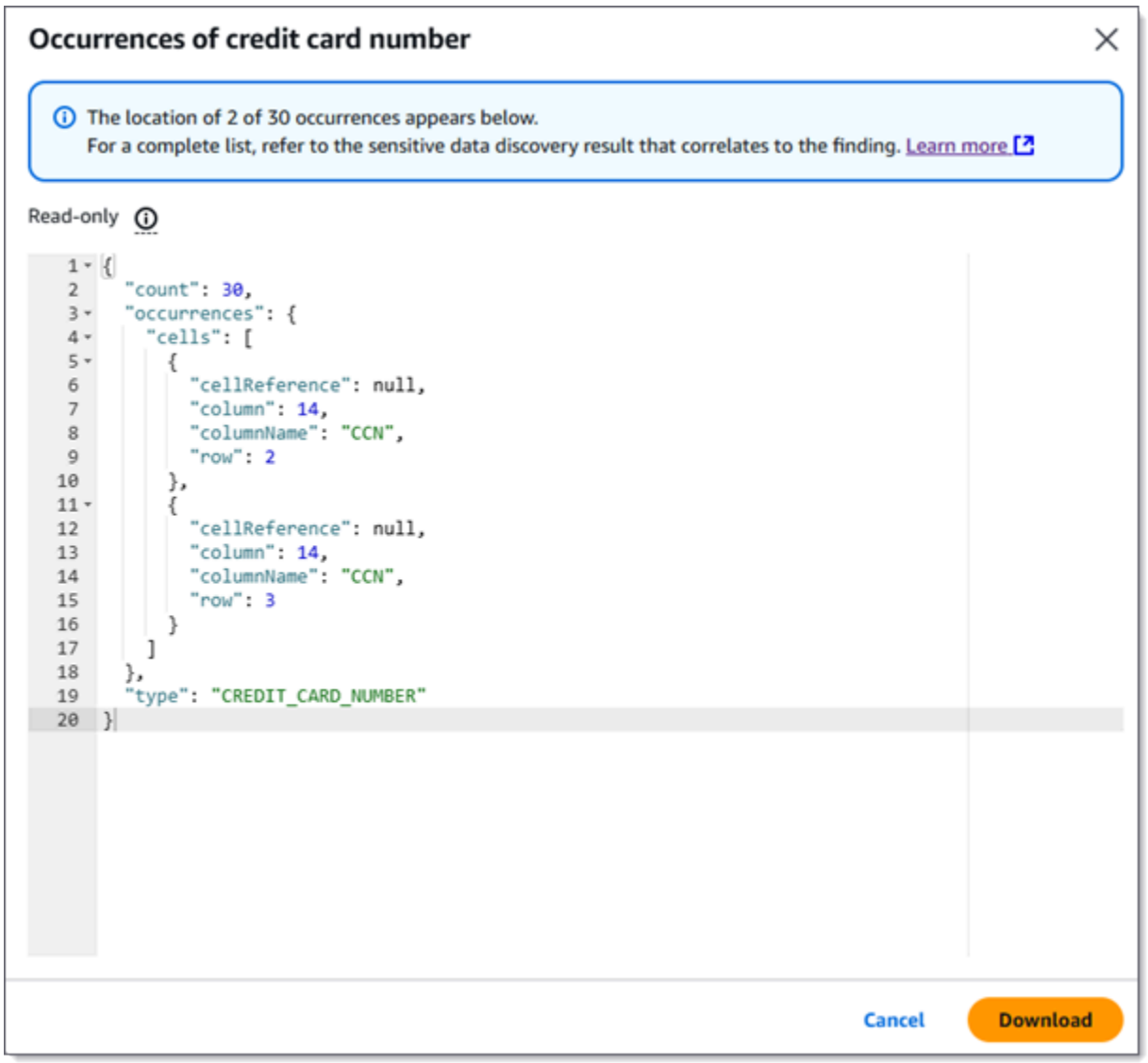
4. In the details panel, scroll to the **Sensitive data** section. This section provides information about the categories and types of sensitive data that Macie found in the affected S3 object. It also indicates the number of occurrences of each type of sensitive data that Macie found.

   For example, the following image shows some details of a finding that reports 30 occurrences of credit card numbers, 20 occurrences of names, and 29 occurrences of US Social Security numbers.



   If the finding includes details about the location of one or more occurrences of a specific type of sensitive data, the number of occurrences is a link. Choose the link to show the details. Macie opens a new window and displays the details in JSON format.

   For example, the following image shows the location of two occurrences of credit card numbers in an affected S3 object.



To save the details as a JSON file, choose **Download**, and then and specify a name and location for the file.

5. To save all the finding's details as a JSON file, choose the finding's identifier (**Finding ID**) at the top of the details panel. Macie opens a new window and displays all the details in JSON format. Choose **Download**, and then specify a name and location for the file.

To access details about the location of as many as 1,000 occurrences of each type of sensitive data in the affected object, refer to the corresponding sensitive data discovery result for the finding. To do this, scroll to the beginning of the **Details** section of the panel. Then choose the link in the **Detailed result location** field. Macie opens the Amazon S3 console and displays the file or folder that contains the corresponding discovery result.

Retrieving sensitive data samples with Macie findings

To verify the nature of sensitive data that Amazon Macie reports in findings, you can optionally configure and use Macie to retrieve and reveal samples of sensitive data reported by individual findings. This includes sensitive data that Macie detects using managed data identifiers, and data that matches the criteria of custom data identifiers. The samples can help you tailor your investigation of an affected Amazon Simple Storage Service (Amazon S3) object and bucket.

If you retrieve and reveal sensitive data samples for a finding, Macie performs the following general tasks:

1. Verifies that the finding specifies the location of individual occurrences of sensitive data and the location of a corresponding sensitive data discovery result.
2. Evaluates the corresponding sensitive data discovery result, checking the validity of the metadata for the affected S3 object and the location data for occurrences of sensitive data in the object.
3. By using data in the sensitive data discovery result, locates the first 1–10 occurrences of sensitive data reported by the finding, and extracts the first 1–128 characters of each occurrence from the affected S3 object. If the finding reports multiple types of sensitive data, Macie does this for up to 100 types.
4. Encrypts the extracted data with an AWS Key Management Service (AWS KMS) key that you specify.
5. Temporarily stores the encrypted data in a cache and displays the data for you to review. The data is encrypted at all times, both in transit and at rest.
6. Soon after extraction and encryption, permanently deletes the data from the cache unless additional retention is temporarily required to resolve an operational issue.

If you choose to retrieve and reveal sensitive data samples for a finding again, Macie repeats these tasks to locate, extract, encrypt, store, and ultimately delete the samples.

Macie doesn't use the Macie service-linked role for your account to perform these tasks. Instead, you use your AWS Identity and Access Management (IAM) identity or allow Macie to assume an IAM role in your account. You can retrieve and reveal sensitive data samples for a finding if you or the role is allowed to access the requisite resources and data, and perform the requisite actions. All the requisite actions are logged in AWS CloudTrail.

Important

We recommend that you restrict access to this functionality by using custom IAM policies. For additional access control, we recommend that you also create a dedicated AWS KMS key for encryption of sensitive data samples that are retrieved, and restrict use of the key to only those principals who must be allowed to retrieve and reveal sensitive data samples. For recommendations and examples of policies that you might use to control access to this functionality, see the following blog post on the *AWS Security Blog*: How to use Amazon Macie to preview sensitive data in S3 buckets.

The topics in this section explain how to configure and use Macie to retrieve and reveal sensitive data samples for findings. You can perform these tasks in all the AWS Regions where Macie is currently available except the Asia Pacific (Osaka) and Israel (Tel Aviv) Regions.

Topics

Configuration options for retrieving sensitive data samples with Macie

You can optionally configure and use Amazon Macie to retrieve and reveal samples of sensitive data that Macie reports in individual findings. If you retrieve and reveal sensitive data samples for a finding, Macie uses data in the corresponding sensitive data discovery result to locate occurrences of sensitive data in the affected Amazon Simple Storage Service (Amazon S3) object. Macie then extracts samples of those occurrences from the affected object. Macie encrypts the extracted data with an AWS Key Management Service (AWS KMS) key that you specify, temporarily stores the encrypted data in a cache, and returns the data in your results for the finding. Soon after extraction and encryption, Macie permanently deletes the data from the cache unless additional retention is temporarily required to resolve an operational issue.

Macie doesn't use the Macie service-linked role for your account to locate, retrieve, encrypt, or reveal sensitive data samples for affected S3 objects. Instead, Macie uses settings and resources that you configure for your account. When you configure the settings in Macie, you specify how to access affected S3 objects. You also specify which AWS KMS key to use to encrypt the samples. You

can configure the settings in all the AWS Regions where Macie is currently available except the Asia Pacific (Osaka) and Israel (Tel Aviv) Regions.

To access affected S3 objects and retrieve sensitive data samples from them, you have two options. You can configure Macie to use AWS Identity and Access Management (IAM) user credentials or assume an IAM role:

- **Use IAM user credentials** – With this option, each user of your account uses their individual IAM identity to locate, retrieve, encrypt, and reveal the samples. This means that a user can retrieve and reveal sensitive data samples for a finding if they're allowed to access the requisite resources and data, and perform the requisite actions.
- **Assume an IAM role** – With this option, you create an IAM role that delegates access to Macie. You also ensure that the trust and permissions policies for the role meet all requirements for Macie to assume the role. Macie then assumes the role when a user of your account chooses to locate, retrieve, encrypt, and reveal sensitive data samples for a finding.

You can use either configuration with any type of Macie account—the delegated Macie administrator account for an organization, a Macie member account in an organization, or a standalone Macie account.

The following topics explain options, requirements, and considerations that can help you determine how to configure the settings and resources for your account. This includes the trust and permissions policies to attach to an IAM role. For additional recommendations and examples of policies that you might use to retrieve and reveal sensitive data samples, see the following blog post on the *AWS Security Blog*: How to use Amazon Macie to preview sensitive data in S3 buckets.

Topics

Determining which access method to use

When determining which configuration is best for your AWS environment, a key consideration is whether your environment includes multiple Amazon Macie accounts that are centrally managed

as an organization. If you're the delegated Macie administrator for an organization, configuring Macie to assume an IAM role can streamline retrieval of sensitive data samples from affected S3 objects for accounts in your organization. With this approach, you create an IAM role in your administrator account. You also create an IAM role in each applicable member account. The role in your administrator account delegates access to Macie. The role in a member account delegates cross-account access to the role in your administrator account. If implemented, you can then use role chaining to access affected S3 objects for your member accounts.

Also consider who has direct access to individual findings by default. To retrieve and reveal sensitive data samples for a finding, a user must first have access to the finding:

- **Sensitive data discovery jobs** – Only the account that creates a job can access findings that the job produces. If you have a Macie administrator account, you can configure a job to analyze objects in S3 buckets for any account in your organization. Therefore, your jobs can produce findings for objects in buckets that your member accounts own. If you have a member account or a standalone Macie account, you can configure a job to analyze objects only in buckets that your account owns.
- **Automated sensitive data discovery** – Only the Macie administrator account can access findings that automated discovery produces for accounts in their organization. Member accounts can't access these findings. If you have a standalone Macie account, you can access findings that automated discovery produces only for your own account.

If you plan to access affected S3 objects by using an IAM role, also consider the following:

- To locate occurrences of sensitive data in an object, the corresponding sensitive data discovery result for a finding must be stored in an S3 object that Macie signed with a Hash-based Message Authentication Code (HMAC) AWS KMS key. Macie must be able to verify the integrity and authenticity of the sensitive data discovery result. Otherwise, Macie doesn't assume the IAM role to retrieve sensitive data samples. This is an additional guardrail for restricting access to data in S3 objects for an account.
- To retrieve sensitive data samples from an object that's encrypted with a customer managed AWS KMS key, the IAM role must be allowed to decrypt data with the key. More specifically, the key's policy must allow the role to perform the `kms:Decrypt` action. For other types of server-side encryption, no additional permissions or resources are required to decrypt an affected object. For more information, see Decrypting affected S3 objects.
- To retrieve sensitive data samples from an object for another account, you must currently be the delegated Macie administrator for the account in the applicable AWS Region. In addition:

- Macie must currently be enabled for the member account in the applicable Region.
- The member account must have an IAM role that delegates cross-account access to an IAM role in your Macie administrator account. The name of the role must be the same in your Macie administrator account and the member account.
- The trust policy for the IAM role in the member account must include a condition that specifies the correct external ID for your configuration. This ID is a unique alphanumeric string that Macie generates automatically after you configure the settings for your Macie administrator account. For information about using external IDs in trust policies, see Access to AWS accounts owned by third parties in the *AWS Identity and Access Management User Guide*.
- If the IAM role in the member account meets all Macie requirements, the member account doesn't need to configure and enable Macie settings for you to retrieve sensitive data samples from objects for their account. Macie uses only the settings and IAM role in your Macie administrator account and the IAM role in the member account.

> **Tip**
>
> If your account is part of a large organization, consider using an AWS CloudFormation template and stack set to provision and manage the IAM roles for member accounts in your organization. For information about creating and using templates and stack sets, see the AWS CloudFormation User Guide.
> To review and optionally download a CloudFormation template that can serve as a starting point, you can use the Amazon Macie console. In the navigation pane on the console, under **Settings**, choose **Reveal samples**. Choose **Edit**, and then choose **View member role permissions and CloudFormation template**.

Subsequent topics in this section provide additional details and considerations for each type of configuration. For IAM roles, this includes the trust and permissions policies to attach to a role. If you're not sure which type of configuration is best for your environment, ask your AWS administrator for assistance.

Using IAM user credentials to access affected S3 objects

If you configure Amazon Macie to retrieve sensitive data samples by using IAM user credentials, each user of your Macie account uses their IAM identity to locate, retrieve, encrypt, and reveal samples for individual findings. This means that a user can retrieve and reveal sensitive data

samples for a finding if their IAM identity is allowed to access the requisite resources and data, and perform the requisite actions. All the requisite actions are logged in AWS CloudTrail.

To retrieve and reveal sensitive data samples for a particular finding, a user must be allowed to access the following data and resources: the finding, the corresponding sensitive data discovery result, the affected S3 bucket, and the affected S3 object. They must also be allowed to use the AWS KMS key that was used to encrypt the affected object, if applicable, and the AWS KMS key that you configure Macie to use to encrypt sensitive data samples. If any IAM policies, resource policies, or other permissions settings deny the requisite access, the user won't be able to retrieve and reveal samples for the finding.

To set up this type of configuration, complete the following general tasks:

1. Verify that you configured a repository for your sensitive data discovery results.
2. Configure the AWS KMS key to use for encryption of sensitive data samples.
3. Verify your permissions for configuring the settings in Macie.
4. Configure and enable the settings in Macie.

For information about performing these tasks, see Configuring Macie to retrieve sensitive data samples.

Assuming an IAM role to access affected S3 objects

To configure Amazon Macie to retrieve sensitive data samples by assuming an IAM role, start by creating an IAM role that delegates access to Amazon Macie. Ensure that the trust and permissions policies for the role meet all requirements for Macie to assume the role. When a user of your Macie account then chooses to retrieve and reveal sensitive data samples for a finding, Macie assumes the role to retrieve the samples from the affected S3 object. Macie assumes the role only when a user chooses to retrieve and reveal samples for a finding. To assume the role, Macie uses the AssumeRole operation of the AWS Security Token Service (AWS STS) API. All the requisite actions are logged in AWS CloudTrail.

To retrieve and reveal sensitive data samples for a particular finding, a user must be allowed to access the finding, the corresponding sensitive data discovery result, and the AWS KMS key that you configure Macie to use to encrypt sensitive data samples. The IAM role must allow Macie to access the affected S3 bucket and the affected S3 object. The role must also be allowed to use the AWS KMS key that was used to encrypt the affected object, if applicable. If any IAM policies,

resource policies, or other permissions settings deny the requisite access, the user won't be able to retrieve and reveal samples for the finding.

To set up this type of configuration, complete the following general tasks. If you have a member account in an organization, work with your Macie administrator to determine whether and how to configure the settings and resources for your account.

1. Define the following:
   - The name of the IAM role that you want Macie to assume. If your account is part of an organization, this name must be same for the delegated Macie administrator account and each applicable member account in the organization. Otherwise, the Macie administrator won't be able to access affected S3 objects for an applicable member account.
   - The name of the IAM permissions policy to attach to the IAM role. If your account is part of an organization, we recommend that you use the same policy name for each applicable member account in the organization. This can streamline provisioning and managing the role in member accounts.
2. Verify that you configured a repository for your sensitive data discovery results.
3. Configure the AWS KMS key to use for encryption of sensitive data samples.
4. Verify your permissions for creating IAM roles and configuring the settings in Macie.
5. If you're the delegated Macie administrator for an organization or you have a standalone Macie account:
   a. Create and configure the IAM role for your account. Ensure that the trust and permissions policies for the role meet all requirements for Macie to assume the role. For details about these requirements, see the next topic.
   b. Configure and enable the settings in Macie. Macie then generates an external ID for the configuration. If you're the Macie administrator for an organization, note this ID. The trust policy for the IAM role in each of your applicable member accounts must specify this ID.
6. If you have a member account in an organization:
   a. Ask your Macie administrator for the external ID to specify in the trust policy for the IAM role in your account. Also verify the name of the IAM role and permissions policy to create.
   b. Create and configure the IAM role for your account. Ensure that the trust and permissions policies for the role meet all requirements for your Macie administrator to assume the role. For details about these requirements, see the next topic.
   c. (Optional) If you want to retrieve and reveal sensitive data samples from affected S3 objects for your own account, configure and enable the settings in Macie. If you want Macie to

assume an IAM role to retrieve the samples, start by creating and configuring an additional IAM role in your account. Ensure that the trust and permissions policies for this additional role meet all requirements for Macie to assume the role. Then configure the settings in Macie and specify the name of this additional role. For details about the policy requirements for the role, see the next topic.

For information about performing these tasks, see Configuring Macie to retrieve sensitive data samples.

Configuring an IAM role to access affected S3 objects

To access affected S3 objects by using an IAM role, start by creating and configuring a role that delegates access to Amazon Macie. Ensure that the trust and permissions policies for the role meet all requirements for Macie to assume the role. How you do this depends on the type of Macie account that you have.

The following sections provide details about the trust and permissions policies to attach to the IAM role for each type of Macie account. Choose the section for the type of account that you have.

Note

If you have a member account in an organization, you might need to create and configure two IAM roles for your account:

- To allow your Macie administrator to retrieve and reveal sensitive data samples from affected S3 objects for your account, create and configure a role that your administrator's account can assume. For these details, choose the **Macie member account** section.
- To retrieve and reveal sensitive data samples from affected S3 objects for your own account, create and configure a role that Macie can assume. For these details, choose the **Standalone Macie account** section.

Before you create and configure either IAM role, work with your Macie administrator to determine the appropriate configuration for your account.

For detailed information about using IAM to create the role, see Creating a role using custom trust policies in the *AWS Identity and Access Management User Guide*.

Macie administrator account

If you're the delegated Macie administrator for an organization, start by using the IAM policy editor to create the permissions policy for the IAM role. The policy should be as follows.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "RetrieveS3Objects",
            "Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "*"
            ]
        },
        {
            "Sid": "AssumeMacieRevealRoleForCrossAccountAccess",
            "Effect": "Allow",
            "Action": [
                "sts:AssumeRole"
            ],
            "Resource": "arn:aws:iam::*:role/IAMRoleName"
        }
    ]
}
```

Where *IAMRoleName* is the name of the IAM role for Macie to assume when retrieving sensitive data samples from affected S3 objects for your organization's accounts. Replace this value with the name of the role that you're creating for your account, and planning to create for applicable member accounts in your organization. This name must be the same for your Macie administrator account and each applicable member account.

Note

In the preceding permissions policy, the `Resource` element in the first statement uses a wildcard character (*). This allows an attached IAM entity to retrieve objects from all the S3 buckets that your organization owns. To allow this access only for specific buckets, replace the wildcard character with the Amazon Resource Name (ARN) of each bucket. For example, to allow access only to objects in a bucket named *amzn-s3-demo-bucket1*, change the element to:
`"Resource": "arn:aws:s3:::amzn-s3-demo-bucket1/*"`
You can also restrict access to objects in specific S3 buckets for individual accounts. To do this, specify bucket ARNs in the `Resource` element of the permissions policy for the IAM role in each applicable account. For more information and examples, see IAM JSON policy elements: Resource in the *AWS Identity and Access Management User Guide*.

After you create the permissions policy for the IAM role, create and configure the role. If you do this by using the IAM console, choose **Custom trust policy** as the **Trusted entity type** for the role. For the trust policy that defines trusted entities for the role, specify the following.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowMacieReveal",
            "Effect": "Allow",
            "Principal": {
                "Service": "reveal-samples.macie.amazonaws.com"
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "111122223333"
                }
            }
        }
    ]
}
```

Where *111122223333* is the account ID for your AWS account. Replace this value with your 12-digit account ID.

In the preceding trust policy:

- The `Principal` element specifies the service principal that Macie uses when retrieving sensitive data samples from affected S3 objects, `reveal-samples.macie.amazonaws.com`.
- The `Action` element specifies the action that the service principal is allowed to perform, the AssumeRole operation of the AWS Security Token Service (AWS STS) API.
- The `Condition` element defines a condition that uses the aws:SourceAccount global condition context key. This condition determines which account can perform the specified action. In this case, it allows Macie to assume the role only for the specified account. The condition helps prevent Macie from being used as a confused deputy during transactions with AWS STS.

After you define the trust policy for the IAM role, attach the permissions policy to the role. This should be the permissions policy that you created before you started creating the role. Then complete the remaining steps in IAM to finish creating and configuring the role. When you finish, configure and enable the settings in Macie.

Macie member account

If you have a Macie member account and you want to allow your Macie administrator to retrieve and reveal sensitive data samples from affected S3 objects for your account, start by asking your Macie administrator for the following information:

- The name of the IAM role to create. The name must be same for your account and the Macie administrator account for your organization.
- The name of the IAM permissions policy to attach to the role.
- The external ID to specify in the trust policy for the role. This ID must be the external ID that Macie generated for your Macie administrator's configuration.

After you receive this information, use the IAM policy editor to create the permissions policy for the role. The policy should be as follows.

JSON

```
{
```

```
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "RetrieveS3Objects",
            "Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "*"
            ]
        }
    ]
}
```

The preceding permissions policy allows an attached IAM entity to retrieve objects from all the S3 buckets for your account. This is because the `Resource` element in the policy uses a wildcard character (*). To allow this access only for specific buckets, replace the wildcard character with the Amazon Resource Name (ARN) of each bucket. For example, to allow access only to objects in a bucket named *amzn-s3-demo-bucket2*, change the element to:

`"Resource": "arn:aws:s3:::amzn-s3-demo-bucket2/*"`

For more information and examples, see IAM JSON policy elements: Resource in the *AWS Identity and Access Management User Guide*.

After you create the permissions policy for the IAM role, create the role. If you create the role by using the IAM console, choose **Custom trust policy** as the **Trusted entity type** for the role. For the trust policy that defines trusted entities for the role, specify the following.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowMacieAdminRevealRoleForCrossAccountAccess",
            "Effect": "Allow",
            "Principal": {
                "AWS": "arn:aws:iam::111122223333:role/IAMRoleName"
            },
```

```
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEquals": {
                    "sts:ExternalId": "externalID",
                    "aws:PrincipalOrgID": "${aws:ResourceOrgID}"
                }
            }
        }
    ]
}
```

In the preceding policy, replace the placeholder values with the correct values for your AWS environment, where:

- *111122223333* is the 12-digit account ID for your Macie administrator's account.
- *IAMRoleName* is the name of the IAM role in your Macie administrator's account. It should be the name that you received from your Macie administrator.
- *externalID* is the external ID that you received from your Macie administrator.

In general, the trust policy allows your Macie administrator to assume the role to retrieve and reveal sensitive data samples from affected S3 objects for your account. The `Principal` element specifies the ARN of an IAM role in your Macie administrator's account. This is the role that your Macie administrator uses to retrieve and reveal sensitive data samples for your organization's accounts. The `Condition` block defines two conditions that further determine who can assume the role:

- The first condition specifies an external ID that's unique to your organization's configuration. To learn more about external IDs, see Access to AWS accounts owned by third parties in the *AWS Identity and Access Management User Guide*.
- The second condition uses the aws:PrincipalOrgID global condition context key. The value for the key is a dynamic variable that represents the unique identifier for an organization in AWS Organizations (`${aws:ResourceOrgID}`). The condition restricts access to only those accounts that are part of the same organization in AWS Organizations. If you joined your organization by accepting an invitation in Macie, remove this condition from the policy.

After you define the trust policy for the IAM role, attach the permissions policy to the role. This should be the permissions policy that you created before you started creating the role. Then

complete the remaining steps in IAM to finish creating and configuring the role. Do not configure and enter settings for the role in Macie.

Standalone Macie account

If you have a standalone Macie account or a Macie member account and you want to retrieve and reveal sensitive data samples from affected S3 objects for your own account, start by using the IAM policy editor to create the permissions policy for the IAM role. The policy should be as follows.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "RetrieveS3Objects",
            "Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "*"
            ]
        }
    ]
}
```

In the preceding permissions policy, the `Resource` element uses a wildcard character (*). This allows an attached IAM entity to retrieve objects from all the S3 buckets for your account. To allow this access only for specific buckets, replace the wildcard character with the Amazon Resource Name (ARN) of each bucket. For example, to allow access only to objects in a bucket named *amzn-s3-demo-bucket3*, change the element to:

```
"Resource": "arn:aws:s3:::amzn-s3-demo-bucket3/*"
```

For more information and examples, see IAM JSON policy elements: Resource in the *AWS Identity and Access Management User Guide*.

After you create the permissions policy for the IAM role, create the role. If you create the role by using the IAM console, choose **Custom trust policy** as the **Trusted entity type** for the role. For the trust policy that defines trusted entities for the role, specify the following.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowMacieReveal",
            "Effect": "Allow",
            "Principal": {
                "Service": "reveal-samples.macie.amazonaws.com"
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "999999999999"
                }
            }
        }
    ]
}
```

Where *999999999999* is the account ID for your AWS account. Replace this value with your 12-digit account ID.

In the preceding trust policy:

- The `Principal` element specifies the service principal that Macie uses when retrieving and revealing sensitive data samples from affected S3 objects, `reveal-samples.macie.amazonaws.com`.
- The `Action` element specifies the action that the service principal is allowed to perform, the AssumeRole operation of the AWS Security Token Service (AWS STS) API.
- The `Condition` element defines a condition that uses the aws:SourceAccount global condition context key. This condition determines which account can perform the specified action. It allows Macie to assume the role only for the specified account. The condition helps prevent Macie from being used as a confused deputy during transactions with AWS STS.

After you define the trust policy for the IAM role, attach the permissions policy to the role. This should be the permissions policy that you created before you started creating the role. Then complete the remaining steps in IAM to finish creating and configuring the role. When you finish, configure and enable the settings in Macie.

Decrypting affected S3 objects

Amazon S3 supports multiple encryption options for S3 objects. For most of these options, no additional resources or permissions are required for an IAM user or role to decrypt and retrieve sensitive data samples from an affected object. This is the case for an object that's encrypted using server-side encryption with an Amazon S3 managed key or an AWS managed AWS KMS key.

However, if an S3 object is encrypted with a customer managed AWS KMS key, additional permissions are required to decrypt and retrieve sensitive data samples from the object. More specifically, the key policy for the KMS key must allow the IAM user or role to perform the `kms:Decrypt` action. Otherwise, an error occurs and Amazon Macie doesn't retrieve any samples from the object. To learn how to provide this access for an IAM user, see KMS key access and permissions in the *AWS Key Management Service Developer Guide*.

How to provide this access for an IAM role depends on whether the account that owns the AWS KMS key also owns the role:

- If the same account owns the KMS key and the role, a user of the account has to update the key's policy.
- If one account owns the KMS key and a different account owns the role, a user of the account that owns the key has to allow cross-account access to the key.

This topic describes how to perform these tasks for an IAM role that you created to retrieve sensitive data samples from S3 objects. It also provides examples for both scenarios. For information about allowing access to customer managed AWS KMS keys for other scenarios, see KMS key access and permissions in the *AWS Key Management Service Developer Guide*.

Allowing same-account access to a customer managed key

If the same account owns both the AWS KMS key and the IAM role, a user of the account has to add a statement to the key's policy. The additional statement must allow the IAM role to decrypt data by using the key. For detailed information about updating a key policy, see Changing a key policy in the *AWS Key Management Service Developer Guide*.

In the statement:

- The `Principal` element must specify the Amazon Resource Name (ARN) of the IAM role.
- The `Action` array must specify the `kms:Decrypt` action. This is the only AWS KMS action that the IAM role must be allowed to perform to decrypt an object that's encrypted with the key.

The following is an example of the statement to add to the policy for a KMS key.

```
{
    "Sid": "Allow the Macie reveal role to use the key",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::123456789012:role/IAMRoleName"
    },
    "Action": [
        "kms:Decrypt"
    ],
    "Resource": "*"
}
```

In the preceding example:

- The `AWS` field in the `Principal` element specifies the ARN of the IAM role in the account. It allows the role to perform the action specified by the policy statement. *123456789012* is an example account ID. Replace this value with the account ID for the account that owns the role and the KMS key. *IAMRoleName* is an example name. Replace this value with the name of the IAM role in the account.
- The `Action` array specifies the action that the IAM role is allowed to perform using the KMS key —decrypt ciphertext that's encrypted with the key.

Where you add this statement to a key policy depends on the structure and elements that the policy currently contains. When you add the statement, ensure that the syntax is valid. Key policies use JSON format. This means that you have to also add a comma before or after the statement, depending on where you add the statement to the policy.

Allowing cross-account access to a customer managed key

If one account owns the AWS KMS key (*key owner*) and a different account owns the IAM role (*role owner*), the key owner has to provide the role owner with cross-account access to the key. One way to do this is by using a grant. A *grant* is a policy instrument that allows AWS principals to use KMS

keys in cryptographic operations if the conditions specified by the grant are met. To learn about grants, see Grants in AWS KMS in the *AWS Key Management Service Developer Guide*.

With this approach, the key owner first ensures that the key's policy allows the role owner to create a grant for the key. The role owner then creates a grant for the key. The grant delegates the relevant permissions to the IAM role in their account. It allows the role to decrypt S3 objects that are encrypted with the key.

Step 1: Update the key policy

In the key policy, the key owner should ensure that the policy includes a statement that allows the role owner to create a grant for the IAM role in their (the role owner's) account. In this statement, the `Principal` element must specify the ARN of the role owner's account. The `Action` array must specify the `kms:CreateGrant` action. A `Condition` block can filter access to the specified action. The following is an example of this statement in the policy for a KMS key.

```
{
    "Sid": "Allow a role in an account to create a grant",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::111122223333:root"
    },
    "Action": [
        "kms:CreateGrant"
    ],
    "Resource": "*",
    "Condition": {
        "StringEquals": {
            "kms:GranteePrincipal": "arn:aws:iam::111122223333:role/IAMRoleName"
        },
        "ForAllValues:StringEquals": {
            "kms:GrantOperations": "Decrypt"
        }
    }
}
```

In the preceding example:

- The `AWS` field in the `Principal` element specifies the ARN of the role owner's account. It allows the account to perform the action specified by the policy statement. *111122223333* is an example account ID. Replace this value with the account ID for the role owner's account.

- The `Action` array specifies the action that the role owner is allowed to perform on the KMS key—create a grant for the key.
- The `Condition` block uses condition operators and the following condition keys to filter access to the action that the role owner is allowed to perform on the KMS key:
  - kms:GranteePrincipal – This condition allows the role owner to create a grant only for the specified grantee principal, which is the ARN of the IAM role in their account. In that ARN, *111122223333* is an example account ID. Replace this value with the account ID for the role owner's account. *IAMRoleName* is an example name. Replace this value with the name of the IAM role in the role owner's account.
  - kms:GrantOperations – This condition allows the role owner to create a grant only to delegate permission to perform the AWS KMS `Decrypt` action (decrypt ciphertext that's encrypted with the key). It prevents the role owner from creating grants that delegate permissions to perform other actions on the KMS key. The `Decrypt` action is the only AWS KMS action that the IAM role must be allowed to perform to decrypt an object that's encrypted with the key.

Where the key owner adds this statement to the key policy depends on the structure and elements that the policy currently contains. When the key owner adds the statement, they should ensure that the syntax is valid. Key policies use JSON format. This means that the key owner has to also add a comma before or after the statement, depending on where they add the statement to the policy. For detailed information about updating a key policy, see Changing a key policy in the *AWS Key Management Service Developer Guide*.

Step 2: Create a grant

After the key owner updates the key policy as necessary, the role owner creates a grant for the key. The grant delegates the relevant permissions to the IAM role in their (the role owner's) account. Before the role owner creates the grant, they should verify that they're allowed to perform the `kms:CreateGrant` action. This action allows them to add a grant to an existing, customer managed AWS KMS key.

To create the grant, the role owner can use the CreateGrant operation of the AWS Key Management Service API. When the role owner creates the grant, they should specify the following values for the required parameters:

- `KeyId` – The ARN of the KMS key. For cross-account access to a KMS key, this value must be an ARN. It can't be a key ID.

- `GranteePrincipal` – The ARN of the IAM role in their account. This value should be `arn:aws:iam::`*`111122223333`*`:role/`*`IAMRoleName`*, where *`111122223333`* is the account ID for the role owner's account and *`IAMRoleName`* is the name of the role.
- `Operations` – The AWS KMS decrypt action (`Decrypt`). This is the only AWS KMS action that the IAM role must be allowed to perform to decrypt an object that's encrypted with the KMS key.

If the role owner is using the AWS Command Line Interface (AWS CLI), they can run the create-grant command to create the grant. The following example shows how. The example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws kms create-grant ^
--key-id arn:aws:kms:us-east-1:123456789012:key/1234abcd-12ab-34cd-56ef-1234567890ab ^
--grantee-principal arn:aws:iam::111122223333:role/IAMRoleName ^
--operations "Decrypt"
```

Where:

- `key-id` specifies the ARN of the KMS key to apply the grant to.
- `grantee-principal` specifies the ARN of the IAM role that's allowed to perform the action specified by the grant. This value should match the ARN specified by the `kms:GranteePrincipal` condition in the key policy.
- `operations` specifies the action that the grant allows the specified principal to perform—decrypt ciphertext that's encrypted with the key.

If the command runs successfully, you receive output similar to the following.

```
{
    "GrantToken": "<grant token>",
    "GrantId": "1a2b3c4d2f5e69f440bae30eaec9570bb1fb7358824f9ddfa1aa5a0dab1a59b2"
}
```

Where `GrantToken` is a unique, non-secret, variable-length, base64-encoded string that represents the grant that was created, and `GrantId` is the unique identifier for the grant.

Configuring Macie to retrieve sensitive data samples

You can optionally configure and use Amazon Macie to retrieve and reveal samples of sensitive data that Macie reports in individual findings. The samples can help you verify the nature of the

sensitive data that Macie found. They can also help you tailor your investigation of an affected Amazon Simple Storage Service (Amazon S3) object and bucket. You can retrieve and reveal sensitive data samples in all the AWS Regions where Macie is currently available except the Asia Pacific (Osaka) and Israel (Tel Aviv) Regions.

When you retrieve and reveal sensitive data samples for a finding, Macie uses data in the corresponding sensitive data discovery result to locate occurrences of sensitive data in the affected S3 object. Macie then extracts samples of those occurrences from the affected object. Macie encrypts the extracted data with an AWS Key Management Service (AWS KMS) key that you specify, temporarily stores the encrypted data in a cache, and returns the data in your results for the finding. Soon after extraction and encryption, Macie permanently deletes the data from the cache unless additional retention is temporarily required to resolve an operational issue.

To retrieve and reveal sensitive data samples for findings, you first need to configure and enable settings for your Macie account. You also need to configure supporting resources and permissions for your account. The topics in this section guide you through the process of configuring Macie to retrieve and reveal sensitive data samples, and managing the status of the configuration for your account.

Topics

> **Tip**
>
> For recommendations and examples of policies that you might use to control access to this functionality, see the following blog post on the *AWS Security Blog*: How to use Amazon Macie to preview sensitive data in S3 buckets.

Before you begin

Before you configure Amazon Macie to retrieve and reveal sensitive data samples for findings, complete the following tasks to ensure that you have the resources and permissions that you need.

Tasks

- Step 1: Configure a repository for sensitive data discovery results
- Step 2: Determine how to access affected S3 objects
- Step 3: Configure an AWS KMS key
- Step 4: Verify your permissions

These tasks are optional if you've already configured Macie to retrieve and reveal sensitive data samples and only want to change your configuration settings.

Step 1: Configure a repository for sensitive data discovery results

When you retrieve and reveal sensitive data samples for a finding, Macie uses data in the corresponding sensitive data discovery result to locate occurrences of sensitive data in the affected S3 object. Therefore, it's important to verify that you configured a repository for your sensitive data discovery results. Otherwise, Macie won't be able to locate sensitive data samples that you want to retrieve and reveal.

To determine whether you've configured this repository for your account, you can use the Amazon Macie console: choose **Discovery results** (under **Settings**) in the navigation pane. To do this programmatically, use the GetClassificationExportConfiguration operation of the Amazon Macie API. To learn more about sensitive data discovery results and how to configure this repository, see Storing and retaining sensitive data discovery results.

Step 2: Determine how to access affected S3 objects

To access affected S3 objects and retrieve sensitive data samples from them, you have two options. You can configure Macie to use your AWS Identity and Access Management (IAM) user credentials. Or you can configure Macie to assume an IAM role that delegates access to Macie. You can use either configuration with any type of Macie account—the delegated Macie administrator account for an organization, a Macie member account in an organization, or a standalone Macie account. Before you configure the settings in Macie, determine which access method you want to use. For details about the options and requirements for each method, see Configuration options for retrieving samples.

If you plan to use an IAM role, create and configure the role before you configure the settings in Macie. Also ensure that the trust and permissions policies for the role meet all requirements for Macie to assume the role. If your account is part of an organization that centrally manages multiple Macie accounts, work with your Macie administrator to first determine whether and how to configure the role for your account.

Step 3: Configure an AWS KMS key

When you retrieve and reveal sensitive data samples for a finding, Macie encrypts the samples with an AWS Key Management Service (AWS KMS) key that you specify. Therefore, you need to determine which AWS KMS key you want to use to encrypt the samples. The key can be an existing KMS key from your own account, or an existing KMS key that another account owns. If you want to use a key that another account owns, obtain the Amazon Resource Name (ARN) of the key. You'll need to specify this ARN when you enter the configuration settings in Macie.

The KMS key must be a customer managed, symmetric encryption key. It must also be a single-Region key that's enabled in the same AWS Region as your Macie account. The KMS key can be in an external key store. However, the key might then be slower and less reliable than a key that's managed entirely within AWS KMS. If latency or an availability issue prevents Macie from encrypting sensitive data samples that you want to retrieve and reveal, an error occurs and Macie doesn't return any samples for the finding.

In addition, the key policy for the key must allow the appropriate principals (IAM roles, IAM users, or AWS accounts) to perform the following actions:

- `kms:Decrypt`
- `kms:DescribeKey`
- `kms:GenerateDataKey`

> **Important**
>
> As an additional layer of access control, we recommend that you create a dedicated KMS key for encryption of sensitive data samples that are retrieved, and restrict use of the key to only those principals who must be allowed to retrieve and reveal sensitive data samples. If a user isn't allowed to perform the preceding actions for the key, Macie rejects their request to retrieve and reveal sensitive data samples. Macie doesn't return any samples for the finding.

For information about creating and configuring KMS keys, see Create a KMS key in the *AWS Key Management Service Developer Guide*. For information about using key policies to manage access to KMS keys, see Key policies in AWS KMS in the *AWS Key Management Service Developer Guide*.

Step 4: Verify your permissions

Before you configure the settings in Macie, also verify that you have the permissions that you need. To verify your permissions, use AWS Identity and Access Management (IAM) to review the IAM policies that are attached to your IAM identity. Then compare the information in those policies to the following list of actions that you must be allowed to perform.

Amazon Macie

For Macie, verify that you're allowed to perform the following actions:

- `macie2:GetMacieSession`
- `macie2:UpdateRevealConfiguration`

The first action allows you to access your Macie account. The second action allows you to change your configuration settings for retrieving and revealing sensitive data samples. This includes enabling and disabling the configuration for your account.

Optionally verify that you're also allowed to perform the `macie2:GetRevealConfiguration` action. This action allows you to retrieve your current configuration settings and the current status of the configuration for your account.

AWS KMS

If you plan to use the Amazon Macie console to enter the configuration settings, also verify that you're allowed to perform the following AWS Key Management Service (AWS KMS) actions:

- `kms:DescribeKey`
- `kms:ListAliases`

These actions allow you to retrieve information about the AWS KMS keys for your account. You can then choose one of these keys when you enter the settings.

IAM

If you plan to configure Macie to assume an IAM role to retrieve and reveal sensitive data samples, also verify that you're allowed to perform the following IAM action: `iam:PassRole`. This action allows you to pass the role to Macie, which in turn allows Macie to assume the role. When you enter the configuration settings for your account, Macie can also then verify that the role exists in your account and is configured correctly.

If you're not allowed to perform the requisite actions, ask your AWS administrator for assistance.

Configuring and enabling Macie settings

After you verify that you have the resources and permissions that you need, you can configure the settings in Amazon Macie and enable the configuration for your account.

If your account is part of an organization that centrally manages multiple Macie accounts, note the following before you configure or subsequently change the settings for your account:

- If you have a member account, work with your Macie administrator to determine whether and how to configure the settings for your account. Your Macie administrator can help you determine the correct configuration settings for your account.
- If you have a Macie administrator account and you change your settings for accessing affected S3 objects, your changes might affect other accounts and resources for your organization. This depends on whether Macie is currently configured to assume an AWS Identity and Access Management (IAM) role to retrieve sensitive data samples. If it is and you reconfigure Macie to use IAM user credentials, Macie permanently deletes existing settings for the IAM role—the name of the role and the external ID for your configuration. If your organization subsequently chooses to use IAM roles again, you'll need to specify a new external ID in the trust policy for the role in each applicable member account.

For details about the configuration options and requirements for either type of account, see Configuration options for retrieving samples.

To configure the settings in Macie and enable the configuration for your account, you can use the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to configure and enable the settings by using the Amazon Macie console.

To configure and enable Macie settings

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to configure and enable Macie to retrieve and reveal sensitive data samples.
3. In the navigation pane, under **Settings**, choose **Reveal samples**.
4. In the **Settings** section, choose **Edit**.

5. For **Status**, choose **Enable**.
6. Under **Access**, specify the access method and settings that you want to use when retrieving sensitive data samples from affected S3 objects:

   - To use an IAM role that delegates access to Macie, choose **Assume an IAM role**. If you choose this option, Macie retrieves the samples by assuming the IAM role that you created and configured in your AWS account. In the **Role name** box, enter the name of the role.
   - To use the credentials of the IAM user who requests the samples, choose **Use IAM user credentials**. If you choose this option, each user of your account uses their individual IAM identity to retrieve the samples.

7. Under **Encryption**, specify the AWS KMS key that you want to use to encrypt sensitive data samples that are retrieved:

   - To use a KMS key from your own account, choose **Select a key from your account**. Then, in the **AWS KMS key** list, choose the key to use. The list displays existing, symmetric encryption KMS keys for your account.
   - To use a KMS key that another account owns, choose **Enter the ARN of a key from another account**. Then, in the **AWS KMS key ARN** box, enter the Amazon Resource Name (ARN) of the key to use—for example, **`arn:aws:kms:us-east-1:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab`**.

8. When you finish entering the settings, choose **Save**.

Macie tests the settings and verifies that they're correct. If you configured Macie to assume an IAM role, Macie also verifies that the role exists in your account and the trust and permissions policies are configured correctly. If there's an issue, Macie displays a message that describes the issue.

To address an issue with the AWS KMS key, refer to the requirements in the preceding topic and specify a KMS key that meets the requirements. To address an issue with the IAM role, start by verifying that you entered the correct role name. If the name is correct, ensure that the role's policies meet all requirements for Macie to assume the role. For these details, see Configuring an IAM role to access affected S3 objects. After you address any issues, you can save and enable the settings.

> **Note**
>
> If you're the Macie administrator for an organization and you configured Macie to assume an IAM role, Macie generates and displays an external ID after you save the settings for your account. Note this ID. The trust policy for the IAM role in each of your applicable member accounts must specify this ID. Otherwise, you won't be able to retrieve sensitive data samples from S3 objects that the accounts own.

API

To configure and enable the settings programmatically, use the UpdateRevealConfiguration operation of the Amazon Macie API. In your request, specify the appropriate values for the supported parameters:

- For the `retrievalConfiguration` parameters, specify the access method and settings that you want to use when retrieving sensitive data samples from affected S3 objects:
  - To assume an IAM role that delegates access to Macie, specify `ASSUME_ROLE` for the `retrievalMode` parameter and specify the name of the role for the `roleName` parameter. If you specify these settings, Macie retrieves the samples by assuming the IAM role that you created and configured in your AWS account.
  - To use the credentials of the IAM user who requests the samples, specify `CALLER_CREDENTIALS` for the `retrievalMode` parameter. If you specify this setting, each user of your account uses their individual IAM identity to retrieve the samples.

> **Important**
>
> If you don't specify values for these parameters, Macie sets the access method (`retrievalMode`) to `CALLER_CREDENTIALS`. If Macie is currently configured to use an IAM role to retrieve the samples, Macie also permanently deletes the current role name and external ID for your configuration. To keep these settings for an existing configuration, include the `retrievalConfiguration` parameters in your request and specify your current settings for those parameters. To retrieve your current settings, use the GetRevealConfiguration operation or, if you're using the AWS Command Line Interface (AWS CLI), run the get-reveal-configuration command.

- For the `kmsKeyId` parameter, specify the AWS KMS key that you want to use to encrypt sensitive data samples that are retrieved:
  - To use a KMS key from your own account, specify the Amazon Resource Name (ARN), ID, or alias for the key. If you specify an alias, include the `alias/` prefix—for example, `alias/ExampleAlias`.
  - To use a KMS key that another account owns, specify the ARN of the key—for example, `arn:aws:kms:us-east-1:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab`. Or specify the ARN of the alias for the key—for example, `arn:aws:kms:us-east-1:111122223333:alias/ExampleAlias`.
- For the `status` parameter, specify `ENABLED` to enable the configuration for your Macie account.

In your request, also ensure that you specify the AWS Region in which you want to enable and use the configuration.

To configure and enable the settings by using the AWS CLI, run the update-reveal-configuration command and specify the appropriate values for the supported parameters. For example, if you're using the AWS CLI on Microsoft Windows, run the following command:

```
C:\> aws macie2 update-reveal-configuration ^
--region us-east-1 ^
--configuration={\"kmsKeyId\":\"arn:aws:kms:us-east-1:111122223333:alias/
ExampleAlias\",\"status\":\"ENABLED\"} ^
--retrievalConfiguration={\"retrievalMode\":\"ASSUME_ROLE\",\"roleName\":
\"MacieRevealRole\"}
```

Where:

- *us-east-1* is the Region in which to enable and use the configuration. In this example, the US East (N. Virginia) Region.
- *arn:aws:kms:us-east-1:111122223333:alias/ExampleAlias* is the ARN of the alias for the AWS KMS key to use. In this example, the key is owned by another account.
- ENABLED is the status of the configuration.
- *ASSUME_ROLE* is the access method to use. In this example, assume the specified IAM role.

- *MacieRevealRole* is the name of the IAM role for Macie to assume when retrieving sensitive data samples.

The preceding example uses the caret (^) line-continuation character to improve readability.

When you submit your request, Macie tests the settings. If you configured Macie to assume an IAM role, Macie also verifies that the role exists in your account and the trust and permissions policies are configured correctly. If there's an issue, your request fails and Macie returns a message that describes the issue. To address an issue with the AWS KMS key, refer to the requirements in the preceding topic and specify a KMS key that meets the requirements. To address an issue with the IAM role, start by verifying that you specified the correct role name. If the name is correct, ensure that the role's policies meet all requirements for Macie to assume the role. For these details, see Configuring an IAM role to access affected S3 objects. After you address the issue, submit your request again.

If your request succeeds, Macie enables the configuration for your account in the specified Region and you receive output similar to the following.

```
{
  "configuration": {
    "kmsKeyId": "arn:aws:kms:us-east-1:111122223333:alias/ExampleAlias",
    "status": "ENABLED"
  },
  "retrievalConfiguration": {
    "externalId": "o2vee30hs31642lexample",
    "retrievalMode": "ASSUME_ROLE",
    "roleName": "MacieRevealRole"
  }
}
```

Where `kmsKeyId` specifies the AWS KMS key to use to encrypt sensitive data samples that are retrieved, and `status` is the status of the configuration for your Macie account. The `retrievalConfiguration` values specify the access method and settings to use when retrieving the samples.

> **Note**
>
> If you're the Macie administrator for an organization and you configured Macie to assume an IAM role, note the external ID (`externalId`) in the response. The trust

policy for the IAM role in each of your applicable member accounts must specify this ID. Otherwise, you won't be able to retrieve sensitive data samples from affected S3 objects that the accounts own.

To subsequently check the settings or status of the configuration for your account, use the GetRevealConfiguration operation or, for the AWS CLI, run the get-reveal-configuration command.

Disabling Macie settings

You can disable the configuration settings for your Amazon Macie account at any time. If you disable the configuration, Macie retains the setting that specifies which AWS KMS key to use to encrypt sensitive data samples that are retrieved. Macie permanently deletes the Amazon S3 access settings for the configuration.

Warning

When you disable the configuration settings for your Macie account, you also permanently delete current settings that specify how to access affected S3 objects. If Macie is currently configured to access affected objects by assuming an AWS Identity and Access Management (IAM) role, this includes: the name of the role, and the external ID that Macie generated for the configuration. These settings can't be recovered after they're deleted.

To disable the configuration settings for your Macie account, you can use the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to disable the configuration settings for your account by using the Amazon Macie console.

To disable Macie settings

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to disable the configuration settings for your Macie account.

3. In the navigation pane, under **Settings**, choose **Reveal samples**.
4. In the **Settings** section, choose **Edit**.
5. For **Status**, choose **Disable**.
6. Choose **Save**.

API

To disable the configuration settings programmatically, use the UpdateRevealConfiguration operation of the Amazon Macie API. In your request, ensure that you specify the AWS Region in which you want to disable the configuration. For the `status` parameter, specify `DISABLED`.

To disable the configuration settings by using the AWS Command Line Interface (AWS CLI), run the update-reveal-configuration command. Use the `region` parameter to specify the Region in which you want to disable the configuration. For the `status` parameter, specify `DISABLED`. For example, if you're using the AWS CLI on Microsoft Windows, run the following command:

```
C:\> aws macie2 update-reveal-configuration --region us-east-1 --
configuration={\"status\":\"DISABLED\"}
```

Where:

- *`us-east-1`* is the Region in which to disable the configuration. In this example, the US East (N. Virginia) Region.
- `DISABLED` is the new status of the configuration.

If your request succeeds, Macie disables the configuration for your account in the specified Region and you receive output similar to the following.

```
{
    "configuration": {
        "status": "DISABLED"
    }
}
```

Where `status` is the new status of the configuration for your Macie account.

If Macie was configured to assume an IAM role to retrieve sensitive data samples, you can optionally delete the role and the role's permissions policy. Macie doesn't delete these resources when you disable the configuration settings for your account. In addition, Macie doesn't use these resources to perform any other tasks for your account. To delete the role and its permissions policy, you can use the IAM console or the IAM API. For more information, see Deleting roles in the *AWS Identity and Access Management User Guide*.

Retrieving sensitive data samples for a Macie finding

By using Amazon Macie, you can retrieve and reveal samples of sensitive data that Macie reports in individual sensitive data findings. This includes sensitive data that Macie detects using managed data identifiers, and data that matches the criteria of custom data identifiers. The samples can help you verify the nature of the sensitive data that Macie found. They can also help you tailor your investigation of an affected Amazon Simple Storage Service (Amazon S3) object and bucket. You can retrieve and reveal sensitive data samples in all the AWS Regions where Macie is currently available except the Asia Pacific (Osaka) and Israel (Tel Aviv) Regions.

If you retrieve and reveal sensitive data samples for a finding, Macie uses data in the corresponding sensitive data discovery result to locate the first 1–10 occurrences of sensitive data reported by the finding. Macie then extracts the first 1–128 characters of each occurrence from the affected S3 object. If a finding reports multiple types of sensitive data, Macie does this for up to 100 types of sensitive data reported by the finding.

When Macie extracts sensitive data from an affected S3 object, Macie encrypts the data with an AWS Key Management Service (AWS KMS) key that you specify, temporarily stores the encrypted data in a cache, and returns the data in your results for the finding. Soon after extraction and encryption, Macie permanently deletes the data from the cache unless additional retention is temporarily required to resolve an operational issue.

If you choose to retrieve and reveal sensitive data samples for a finding again, Macie repeats the process for locating, extracting, encrypting, storing, and ultimately deleting the samples.

For a demonstration of how you can retrieve and reveal sensitive data samples by using the Amazon Macie console, watch the following video: Retrieving and revealing sensitive data samples with Amazon Macie.

Topics

- Before you begin

- Determining whether sensitive data samples are available for a finding
- Retrieving sensitive data samples for a finding

Before you begin

Before you can retrieve and reveal sensitive data samples for findings, you need to configure and enable settings for your Amazon Macie account. You also need to work with your AWS administrator to verify that you have the permissions and resources that you need.

When you retrieve and reveal sensitive data samples for a finding, Macie performs a series of tasks to locate, retrieve, encrypt, and reveal the samples. Macie doesn't use the Macie service-linked role for your account to perform these tasks. Instead, you use your AWS Identity and Access Management (IAM) identity or allow Macie to assume an IAM role in your account.

To retrieve and reveal sensitive data samples for a finding, you must have access to the finding, the corresponding sensitive data discovery result, and the AWS KMS key that you configured Macie to use to encrypt sensitive data samples. In addition, you or the IAM role must be allowed to access the affected S3 bucket and the affected S3 object. You or the role must also be allowed to use the AWS KMS key that was used to encrypt the affected object, if applicable. If any IAM policies, resource policies, or other permissions settings deny the requisite access, an error occurs and Macie doesn't return any samples for the finding.

You must also be allowed to perform the following Macie actions:

- `macie2:GetMacieSession`
- `macie2:GetFindings`
- `macie2:ListFindings`
- `macie2:GetSensitiveDataOccurrences`

The first three actions allow you to access your Macie account and retrieve the details of findings. The last action allows you to retrieve and reveal sensitive data samples for findings.

To use the Amazon Macie console to retrieve and reveal sensitive data samples, you must also be allowed to perform the following action: `macie2:GetSensitiveDataOccurrencesAvailability`. This action allows you to determine whether samples are available for individual findings. You don't need permission to perform this action to retrieve and reveal samples programmatically. However, having this permission can streamline your retrieval of samples.

If you're the delegated Macie administrator for an organization and you configured Macie to assume an IAM role to retrieve sensitive data samples, you must also be allowed to perform the following action: `macie2:GetMember`. This action allows you to retrieve information about the association between your account and an affected account. It enables Macie to verify that you're currently the Macie administrator for the affected account.

If you're not allowed to perform the requisite actions or access the requisite data and resources, ask your AWS administrator for assistance.

Determining whether sensitive data samples are available for a finding

To retrieve and reveal sensitive data samples for a finding, the finding needs to meet certain criteria. It has to include location data for specific occurrences of sensitive data. In addition, it has to specify the location of a valid, corresponding sensitive data discovery result. The sensitive data discovery result must be stored in the same AWS Region as the finding. If you configured Amazon Macie to access affected S3 objects by assuming an AWS Identity and Access Management (IAM) role, the sensitive data discovery result must also be stored in an S3 object that Macie signed with a Hash-based Message Authentication Code (HMAC) AWS KMS key.

The affected S3 object also needs to meet certain criteria. The MIME type of the object must be one of the following:

- *application/avro*, for an Apache Avro object container (.avro) file
- *application/gzip*, for a GNU Zip compressed archive (.gz or .gzip) file
- *application/json*, for a JSON or JSON Lines (.json or .jsonl) file
- *application/parquet*, for an Apache Parquet (.parquet) file
- *application/vnd.openxmlformats-officedocument.spreadsheetml.sheet*, for a Microsoft Excel workbook (.xlsx) file
- *application/zip*, for a ZIP compressed archive (.zip) file
- *text/csv*, for a CSV (.csv) file
- *text/plain*, for a non-binary text file other than a CSV, JSON, JSON Lines, or TSV file
- *text/tab-separated-values*, for a TSV (.tsv) file

In addition, the contents of the S3 object must be the same as when the finding was created. Macie checks the object's entity tag (ETag) to determine whether it matches the ETag specified by the finding. Also, the storage size of the object can't exceed the applicable size quota for retrieving and revealing sensitive data samples. For a list of applicable quotas, see Quotas for Macie.

If a finding and the affected S3 object meet the preceding criteria, sensitive data samples are available for the finding. You can optionally determine whether this is the case for a particular finding before you try to retrieve and reveal samples for it.

To determine whether sensitive data samples are available for a finding

You can use the Amazon Macie console or the Amazon Macie API to determine whether sensitive data samples are available for a finding.

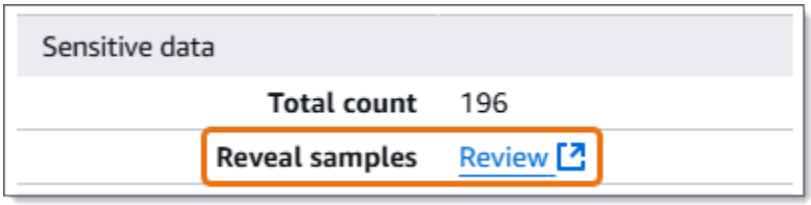
Console

Follow these steps on the Amazon Macie console to determine whether sensitive data samples are available for a finding.

To determine whether samples are available for a finding

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Findings**.
3. On the **Findings** page, choose the finding. The details panel displays information for the finding.
4. In the details panel, scroll to the **Sensitive data** section. Then refer to the **Reveal samples** field.

   If sensitive data samples are available for the finding, a **Review** link appears in the field, as shown in the following image.



   If sensitive data samples aren't available for the finding, the **Reveal samples** field displays text indicating why:

   - **Account not in organization** – You're not allowed to access the affected S3 object by using Macie. The affected account isn't currently part of your organization. Or the account is part of your organization but Macie isn't currently enabled for the account in the current AWS Region.

- **Invalid classification result** – There isn't a corresponding sensitive data discovery result for the finding. Or the corresponding sensitive data discovery result isn't available in the current AWS Region, is malformed or corrupted, or uses an unsupported storage format. Macie can't verify the location of the sensitive data to retrieve.
- **Invalid result signature** – The corresponding sensitive data discovery result is stored in an S3 object that wasn't signed by Macie. Macie can't verify the integrity and authenticity of the sensitive data discovery result. Therefore, Macie can't verify the location of the sensitive data to retrieve.
- **Member role too permissive** – The trust or permissions policy for the IAM role in the affected member account doesn't meet Macie requirements for restricting access to the role. Or the role's trust policy doesn't specify the correct external ID for your organization. Macie can’t assume the role to retrieve the sensitive data.
- **Missing GetMember permission** – You're not allowed to retrieve information about the association between your account and the affected account. Macie can't determine whether you’re allowed to access the affected S3 object as the delegated Macie administrator for the affected account.
- **Object exceeds size quota** – The storage size of the affected S3 object exceeds the size quota for retrieving and revealing samples of sensitive data from that type of file.
- **Object unavailable** – The affected S3 object isn't available. The object was renamed, moved, or deleted, or its contents changed after Macie created the finding. Or the object is encrypted with an AWS KMS key that isn’t available. For example, the key is disabled, is scheduled for deletion, or was deleted.
- **Result not signed** – The corresponding sensitive data discovery result is stored in an S3 object that hasn't been signed. Macie can't verify the integrity and authenticity of the sensitive data discovery result. Therefore, Macie can't verify the location of the sensitive data to retrieve.
- **Role too permissive** – Your account is configured to retrieve occurrences of sensitive data by using an IAM role whose trust or permissions policy doesn't meet Macie requirements for restricting access to the role. Macie can’t assume the role to retrieve the sensitive data.
- **Unsupported object type** – The affected S3 object uses a file or storage format that Macie doesn't support for retrieving and revealing samples of sensitive data. The MIME type of the affected S3 object isn't one of the values in the preceding list.

If there's an issue with the sensitive data discovery result for the finding, the information in the **Detailed result location** field of the finding can help you investigate the issue. This field specifies the original path to the result in Amazon S3. To investigate an issue with an IAM role, ensure that the role's policies meet all requirements for Macie to assume the role. For these details, see Configuring an IAM role to access affected S3 objects.

API

To programmatically determine whether sensitive data samples are available for a finding, use the GetSensitiveDataOccurrencesAvailability operation of the Amazon Macie API. When you submit your request, use the `findingId` parameter to specify the unique identifier for the finding. To obtain this identifier, you can use the ListFindings operation.

If you're using the AWS Command Line Interface (AWS CLI), run the get-sensitive-data-occurrences-availability command and use the `finding-id` parameter to specify the unique identifier for the finding. To obtain this identifier, you can run the list-findings command.

If your request succeeds and samples are available for the finding, you receive output similar to the following:

```
{
    "code": "AVAILABLE",
    "reasons": []
}
```

If your request succeeds and samples aren't available for the finding, the value for the `code` field is `UNAVAILABLE` and the `reasons` array specifies why. For example:

```
{
    "code": "UNAVAILABLE",
    "reasons": [
        "UNSUPPORTED_OBJECT_TYPE"
    ]
}
```

If there's an issue with the sensitive data discovery result for the finding, the information in the `classificationDetails.detailedResultsLocation` field of the finding can help you investigate the issue. This field specifies the original path to the result in Amazon S3. To

investigate an issue with an IAM role, ensure that the role's policies meet all requirements for Macie to assume the role. For these details, see Configuring an IAM role to access affected S3 objects.

Retrieving sensitive data samples for a finding

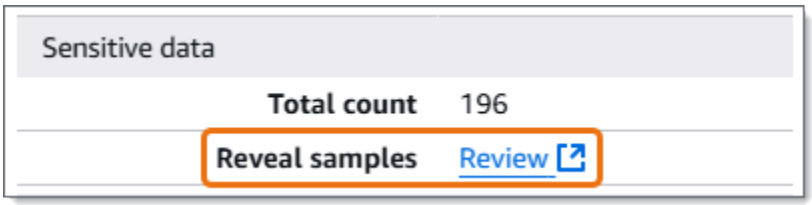
To retrieve and reveal sensitive data samples for a finding, you can use the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to retrieve and reveal sensitive data samples for a finding by using the Amazon Macie console.

To retrieve and reveal sensitive data samples for a finding

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Findings**.
3. On the **Findings** page, choose the finding. The details panel displays information for the finding.
4. In the details panel, scroll to the **Sensitive data** section. Then, in the **Reveal samples** field, choose **Review**:

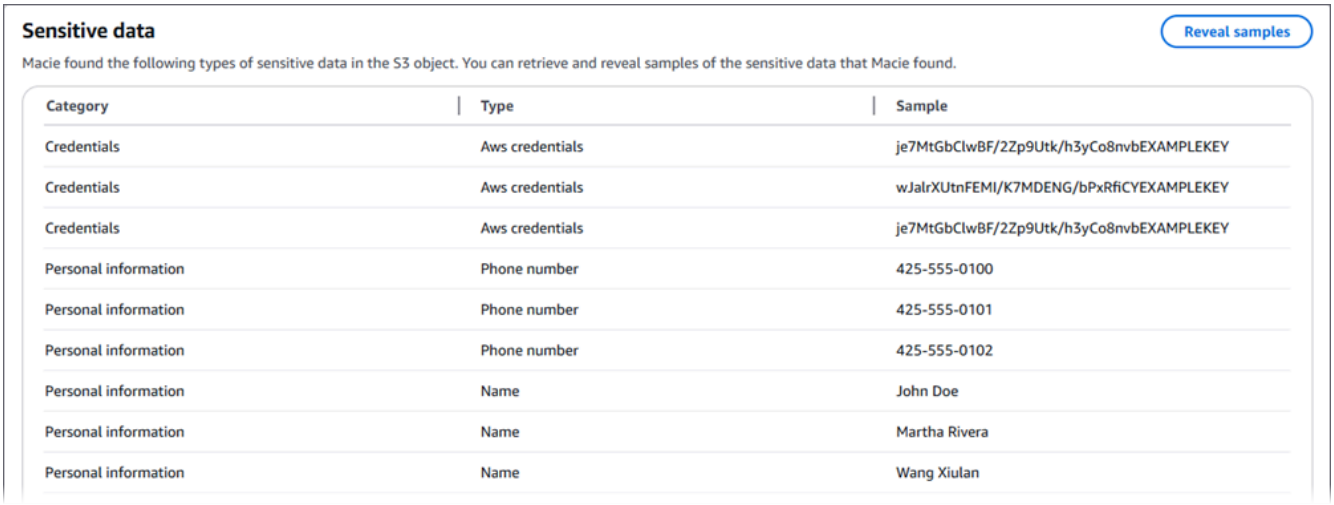


> **Note**
>
> If the **Review** link doesn't appear in the **Reveal samples** field, sensitive data samples aren't available for the finding. To determine why this is the case, see the preceding topic.

After you choose **Review**, Macie displays a page that summarizes key details of the finding. The details include the categories, types, and number of occurrences of sensitive data that Macie found in the affected S3 object.

5. In the **Sensitive data** section of the page, choose **Reveal samples**. Macie then retrieves and reveals samples of the first 1–10 occurrences of sensitive data reported by the finding. Each sample contains the first 1–128 characters of an occurrence of sensitive data. It can take several minutes to retrieve and reveal the samples.

   If the finding reports multiple types of sensitive data, Macie retrieves and reveals samples for up to 100 types. For example, the following image shows samples that span multiple categories and types of sensitive data—AWS credentials, US phone numbers, and people's names.



   The samples are organized first by sensitive data category, and then by sensitive data type.

API

To retrieve and reveal sensitive data samples for a finding programmatically, use the GetSensitiveDataOccurrences operation of the Amazon Macie API. When you submit your request, use the `findingId` parameter to specify the unique identifier for the finding. To obtain this identifier, you can use the ListFindings operation.

To retrieve and reveal sensitive data samples by using the AWS Command Line Interface (AWS CLI), run the get-sensitive-data-occurrences command and use the `finding-id` parameter to specify the unique identifier for the finding. For example:

```
C:\> aws macie2 get-sensitive-data-occurrences --finding-id
 "1f1c2d74db5d8caa76859ec52example"
```

Where *1f1c2d74db5d8caa76859ec52example* is the unique identifier for the finding. To obtain this identifier by using the AWS CLI, you can run the list-findings command.

If your request succeeds, Macie begins processing your request and you receive output similar to the following:

```
{
    "status": "PROCESSING"
}
```

It can take several minutes to process your request. Within a few minutes, submit your request again.

If Macie can locate, retrieve, and encrypt the sensitive data samples, Macie returns the samples in a `sensitiveDataOccurrences` map. The map specifies 1–100 types of sensitive data reported by the finding and 1–10 samples for each type. Each sample contains the first 1–128 characters of an occurrence of sensitive data reported by the finding.

In the map, each key is the ID of the managed data identifier that detected the sensitive data, or the name and unique identifier for the custom data identifier that detected the sensitive data. The values are samples for the specified managed data identifier or custom data identifier. For example, the following response provides three samples of people's names and two samples of AWS secret access keys that were detected by managed data identifiers (`NAME` and `AWS_CREDENTIALS`, respectively).

```
{
    "sensitiveDataOccurrences": {
        "NAME": [
            {
                "value": "Akua Mansa"
            },
            {
                "value": "John Doe"
            },
            {
                "value": "Martha Rivera"
            }
        ],
        "AWS_CREDENTIALS": [
            {
                "value": "wJalrXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY"
```

```
            },
            {
                "value": "je7MtGbClwBF/2Zp9Utk/h3yCo8nvbEXAMPLEKEY"
            }
        ]
    },
    "status": "SUCCESS"
}
```

If your request succeeds but sensitive data samples aren't available for the finding, you receive an `UnprocessableEntityException` message that indicates why samples aren't available. For example:

```
{
    "message": "An error occurred (UnprocessableEntityException) when calling the
 GetSensitiveDataOccurrences operation: OBJECT_UNAVAILABLE"
}
```

In the preceding example, Macie attempted to retrieve samples from the affected S3 object but the object isn't available anymore. The contents of the object changed after Macie created the finding.

If your request succeeds but another type of error prevented Macie from retrieving and revealing sensitive data samples for the finding, you receive output similar to the following:

```
{
    "error": "Macie can't retrieve the samples. You're not allowed to access the
 affected S3 object or the object is encrypted with a key that you're not allowed to
 use.",
    "status": "ERROR"
}
```

The value for the `status` field is `ERROR` and the `error` field describes the error that occurred. The information in the preceding topic can help you investigate the error.

Schema for reporting the location of sensitive data

Amazon Macie uses standardized JSON structures to store information about where it finds sensitive data in Amazon Simple Storage Service (Amazon S3) objects. The structures are used by sensitive data findings and sensitive data discovery results. For sensitive data findings, the

structures are part of the JSON schema for findings. To review the complete JSON schema for findings, see Findings in the *Amazon Macie API Reference*. To learn more about sensitive data discovery results, see Storing and retaining sensitive data discovery results.

Topics

Schema overview

To report the location of sensitive data that Amazon Macie found in an affected S3 object, the JSON schema for sensitive data findings and sensitive data discovery results includes one `customDataIdentifiers` object and one `sensitiveData` object. The `customDataIdentifiers` object provides details about data that Macie detected using custom data identifiers. The `sensitiveData` object provides details about data that Macie detected using managed data identifiers.

Each `customDataIdentifiers` and `sensitiveData` object contains one or more `detections` arrays:

- In a `customDataIdentifiers` object, the `detections` array indicates which custom data identifiers detected the data and produced the finding. For each custom data identifier, the array also indicates the number of occurrences of the data that the identifier detected. It can also indicate the location of the data that the identifier detected.
- In a `sensitiveData` object, a `detections` array indicates the types of sensitive data that Macie detected using managed data identifiers. For each type of sensitive data, the array also indicates the number of occurrences of the data, and it can indicate the location of the data.

For a sensitive data finding, a `detections` array can include 1–15 `occurrences` objects. Each `occurrences` object specifies where Macie detected individual occurrences of a specific type of sensitive data.

For example, the following `detections` array indicates the location of three occurrences of sensitive data (US Social Security numbers) that Macie found in a CSV file.

```
"sensitiveData": [
     {
```

```
      "category": "PERSONAL_INFORMATION",
      "detections": [
         {
            "count": 30,
            "occurrences": {
               "cells": [
                  {
                     "cellReference": null,
                     "column": 1,
                     "columnName": "SSN",
                     "row": 2
                  },
                  {
                     "cellReference": null,
                     "column": 1,
                     "columnName": "SSN",
                     "row": 3
                  },
                  {
                     "cellReference": null,
                     "column": 1,
                     "columnName": "SSN",
                     "row": 4
                  }
               ]
            },
            "type": "USA_SOCIAL_SECURITY_NUMBER"
         }
```

The location and number of `occurrences` objects in a `detections` array varies based on the categories, types, and number of occurrences of sensitive data that Macie detects during an automated sensitive data discovery analysis cycle or a run of a sensitive data discovery job. For each analysis cycle or job run, Macie uses a *depth-first search* algorithm to populate the resulting findings with location data for 1–15 occurrences of sensitive data that Macie detects in S3 objects. These occurrences are indicative of the categories and types of sensitive data that an affected S3 bucket and object might contain.

An `occurrences` object can contain any the following structures, depending on an affected S3 object's file type or storage format:

- `cells` array – This array applies to Microsoft Excel workbooks, CSV files, and TSV files. An object in this array specifies a cell or field that Macie detected an occurrence of sensitive data in.

- `lineRanges` array – This array applies to email message (EML) files, and non-binary text files other than CSV, JSON, JSON Lines, and TSV files—for example, HTML, TXT, and XML files. An object in this array specifies a line or an inclusive range of lines that Macie detected an occurrence of sensitive data in, and the position of the data on the specified line or lines.

  In certain cases, an object in a `lineRanges` array specifies the location of a sensitive data detection in a file type or storage format that's supported by another type of array. Those cases are: a detection in an unstructured section of an otherwise structured file, such as a comment in a file; a detection in a malformed file that Macie analyzes as plaintext; and, a CSV or TSV file that has one or more column names that Macie detected sensitive data in.
- `offsetRanges` array – This array is reserved for future use. If this array is present, the value for it is null.
- `pages` array – This array applies to Adobe Portable Document Format (PDF) files. An object in this array specifies a page that Macie detected an occurrence of sensitive data in.
- `records` array – This array applies to Apache Avro object containers, Apache Parquet files, JSON files, and JSON Lines files. For Avro object containers and Parquet files, an object in this array specifies a record index and the path to a field in a record that Macie detected an occurrence of sensitive data in. For JSON and JSON Lines files, an object in this array specifies the path to a field or array that Macie detected an occurrence of sensitive data in. For JSON Lines files, it also specifies the index of the line that contains the data.

The contents of these arrays vary based on an affected S3 object's file type or storage format and its contents.

Schema details and examples

Amazon Macie tailors the contents of the JSON structures that it uses to indicate where it detected sensitive data in specific types of files and content. The following topics explain and provide examples of these structures.

Topics

For a complete list of JSON structures that can be included in a sensitive data finding, see Findings in the *Amazon Macie API Reference*.

Cells array

Applies to: Microsoft Excel workbooks, CSV files, and TSV files

In a `cells` array, a `Cell` object specifies a cell or field that Macie detected an occurrence of sensitive data in. The following table describes the purpose of each field in a `Cell` object.

Field	Type	Description
`cellReference`	String	The location of the cell, as an absolute cell reference, that contains the occurrence. This field applies only to Excel workbooks. This value is null for CSV and TSV files.
`column`	Integer	The column number of the column that contains the occurrence. For an Excel workbook, this value correlates to the alphabeti cal character(s) for a column identifier—for example, 1 for column A, 2 for column B, and so on.
`columnName`	String	The name of the column that contains the occurrence, if available.
`row`	Integer	The row number of the row that contains the occurrence.

The following example shows the structure of a `Cell` object that specifies the location of an occurrence of sensitive data that Macie detected in a CSV file.

```
"cells": [
   {
      "cellReference": null,
      "column": 3,
      "columnName": "SSN",
      "row": 5
   }
]
```

In the preceding example, the finding indicates that Macie detected sensitive data in the field in the fifth row of the third column (named *SSN*) of the file.

The following example shows the structure of a `Cell` object that specifies the location of an occurrence of sensitive data that Macie detected in an Excel workbook.

```
"cells": [
   {
      "cellReference": "Sheet2!C5",
      "column": 3,
      "columnName": "SSN",
      "row": 5
   }
]
```

In the preceding example, the finding indicates that Macie detected sensitive data in the worksheet named *Sheet2* in the workbook. In that worksheet, Macie detected sensitive data in the cell in the fifth row of the third column (column C, named *SSN*).

LineRanges array

Applies to: Email message (EML) files, and non-binary text files other than CSV, JSON, JSON Lines, and TSV files—for example, HTML, TXT, and XML files

In a `lineRanges` array, a `Range` object specifies a line or an inclusive range of lines that Macie detected an occurrence of sensitive data in, and the position of the data on the specified line or lines.

This object is often empty for file types that are supported by other types of arrays in `occurrences` objects. Exceptions are:

- Data in unstructured sections of an otherwise structured file, such as a comment in a file.

- Data in a malformed file that Macie analyzes as plaintext.
- A CSV or TSV file that has one or more column names that Macie detected sensitive data in.

The following table describes the purpose of each field in a `Range` object of a `lineRanges` array.

Field	Type	Description
`end`	Integer	The number of lines from the beginning of the file to the end of the occurrence.
`start`	Integer	The number of lines from the beginning of the file to the beginning of the occurrence.
`startColumn`	Integer	The number of characters, with spaces and starting from 1, from the beginning of the first line that contains the occurrence (`start`) to the beginning of the occurrence.

The following example shows the structure of a `Range` object that specifies the location of an occurrence of sensitive data that Macie detected on a single line in a TXT file.

```
"lineRanges": [
   {
      "end": 1,
      "start": 1,
      "startColumn": 119
   }
]
```

In the preceding example, the finding indicates that Macie detected a complete occurrence of sensitive data (a mailing address) in the first line of the file. The first character in the occurrence is 119 characters (with spaces) from the beginning of that line.

The following example shows the structure of a Range object that specifies the location of an occurrence of sensitive data that spans multiple lines in a TXT file.

```
"lineRanges": [
   {
      "end": 54,
      "start": 51,
      "startColumn": 1
   }
]
```

In the preceding example, the finding indicates that Macie detected an occurrence of sensitive data (a mailing address) spanning lines 51 through 54 of the file. The first character in the occurrence is the first character on line 51 of the file.

Pages array

Applies to: Adobe Portable Document Format (PDF) files

In a pages array, a Page object specifies a page that Macie detected an occurrence of sensitive data in. The object contains a pageNumber field. The pageNumber field stores an integer that specifies the page number of the page that contains the occurrence.

The following example shows the structure of a Page object that specifies the location of an occurrence of sensitive data that Macie detected in a PDF file.

```
"pages": [
   {
      "pageNumber": 10
   }
]
```

In the preceding example, the finding indicates that page 10 of the file contains the occurrence.

Records array

Applies to: Apache Avro object containers, Apache Parquet files, JSON files, and JSON Lines files

For an Avro object container or a Parquet file, a Record object in a records array specifies a record index and the path to a field in a record that Macie detected an occurrence of sensitive data

in. For JSON and JSON Lines files, a `Record` object specifies the path to a field or array that Macie detected an occurrence of sensitive data in. For JSON Lines files, it also specifies the index of the line that contains the occurrence.

The following table describes the purpose of each field in a `Record` object.

Field	Type	Description
`jsonPath`	String	The path, as a JSONPath expression, to the occurrence. For an Avro object container or a Parquet file, this is the path to the field in the record (`recordIndex`) that contains the occurrence. For a JSON or JSON Lines file, this is the path to the field or array that contains the occurrence. If the data is a value in an array, the path also indicates which value contains the occurrence. If Macie detects sensitive data in the name of any element in the path, Macie omits the `jsonPath` field from a `Record` object. If the name of a path element exceeds 240 characters, Macie truncates the name by removing characters from the beginning of the name. If the resulting full path exceeds 250 characters, Macie also truncates the path, sta

Field	Type	Description
		rting with the first element in the path, until the path cont ains 250 or fewer characters.
`recordIndex`	Integer	For an Avro object container or a Parquet file, the record index, starting from 0, for the record that contains the occurrence. For a JSON Lines file, the line index, starting from 0, for the line that contains the occurrence. This value is always `0` for JSON files.

The following example shows the structure of a `Record` object that specifies the location of an occurrence of sensitive data that Macie detected in a Parquet file.

```
"records": [
   {
      "jsonPath": "$['abcdefghijklmnopqrstuvwxyz']",
      "recordIndex": 7663
   }
]
```

In the preceding example, the finding indicates that Macie detected sensitive data in the record of index 7663 (record number 7664). In that record, Macie detected sensitive data in the field named `abcdefghijklmnopqrstuvwxyz`. The full JSON path to the field in the record is `$.abcdefghijklmnopqrstuvwxyz`. The field is a direct descendant of the root (outer-level) object.

The following example also shows the structure of a `Record` object for an occurrence of sensitive data that Macie detected in a Parquet file. However, in this example, Macie truncated the name of the field that contains the occurrence because the name exceeds the character limit.

```
"records": [
```

```
   {
      "jsonPath":
 "$['...uvwxyzabcdefghijklmnopqrstuvwxyzabcdefghijklmnopqrstuvwxyzabcdefghijklmnopqrstuvwxyzabc
      "recordIndex": 7663
   }
]
```

In the preceding example, the field is a direct descendant of the root (outer-level) object.

In the following example, also for an occurrence of sensitive data that Macie detected in a Parquet file, Macie truncated the full path to the field that contains the occurrence. The full path exceeds the character limit.

```
"records": [
   {
      "jsonPath":
 "$..usssn2.usssn3.usssn4.usssn5.usssn6.usssn7.usssn8.usssn9.usssn10.usssn11.usssn12.usssn13.us
      "recordIndex": 2335
   }
]
```

In the preceding example, the finding indicates that Macie detected sensitive data in the record of index 2335 (record number 2336). In that record, Macie detected sensitive data in the field named `abcdefghijklmnopqrstuvwxyz`. The full JSON path to the field in the record is:

```
$['1234567890']usssn1.usssn2.usssn3.usssn4.usssn5.usssn6.usssn7.usssn8.usssn9.us
```

The following example shows the structure of a `Record` object that specifies the location of an occurrence of sensitive data that Macie detected in a JSON file. In this example, the occurrence is a specific value in an array.

```
"records": [
   {
      "jsonPath": "$.access.key[2]",
      "recordIndex": 0
   }
]
```

In the preceding example, the finding indicates that Macie detected sensitive data in the second value of an array named `key`. The array is a child of an object named `access`.

The following example shows the structure of a `Record` object that specifies the location of an occurrence of sensitive data that Macie detected in a JSON Lines file.

```
"records": [
   {
      "jsonPath": "$.access.key",
      "recordIndex": 3
   }
]
```

In the preceding example, the finding indicates that Macie detected sensitive data in the third value (line) in the file. In that line, the occurrence is in a field named `key`, which is a child of an object named `access`.

Suppressing Macie findings

To streamline your analysis of findings, you can create and use suppression rules. A *suppression rule* is a set of attribute-based filter criteria that defines cases where you want Amazon Macie to archive findings automatically. Suppression rules are helpful in situations where you've reviewed a class of findings and don't want to be notified of them again.

For example, you might decide to allow S3 buckets to contain mailing addresses, if the buckets don't allow public access and they encrypt new objects automatically with a particular AWS KMS key. In this case, you can create a suppression rule that specifies filter criteria for the following fields: **Sensitive data detection type**, **S3 bucket public access permission**, and **S3 bucket encryption KMS key id**. The rule suppresses future findings that match the filter criteria.

If you suppress findings with a suppression rule, Macie continues to generate findings for subsequent occurrences of sensitive data and potential policy violations that match the rule's criteria. However, Macie automatically changes the status of the findings to *archived*. This means that the findings don't appear by default on the Amazon Macie console, but they persist in Macie until they expire. Macie stores findings for 90 days.

In addition, Macie doesn't publish suppressed findings to Amazon EventBridge as events or to AWS Security Hub. Macie does, however, continue to create and store sensitive data discovery results that correlate to sensitive data findings that you suppress. This helps ensure that you have an immutable history of sensitive data findings for data privacy and protection audits or investigations that you perform.

Note

If your account is part of an organization that centrally manages multiple Macie accounts, suppression rules might work differently for your account. This depends on the category of findings that you want to suppress, and whether you have a Macie administrator or member account:

- **Policy findings** – Only a Macie administrator can suppress policy findings for the organization's accounts.

  If you have a Macie administrator account and you create a suppression rule, Macie applies the rule to policy findings for all the accounts in your organization unless you configure the rule to exclude specific accounts. If you have a member account and you want to suppress policy findings for your account, contact your Macie administrator.

- **Sensitive data findings** – A Macie administrator and individual members can suppress sensitive data findings that their sensitive data discovery jobs produce. A Macie administrator can also suppress findings that Macie generates while performing automated sensitive data discovery for the organization.

  Only the account that creates a sensitive data discovery job can suppress or otherwise access sensitive data findings that the job produces. Only the Macie administrator account for an organization can suppress or otherwise access findings that automated sensitive data discovery produces for accounts in the organization.

For more information about the tasks that administrators and members can perform, see Macie administrator and member account relationships.

Topics

Creating a suppression rule for Macie findings

A *suppression rule* is a set of attribute-based filter criteria that defines cases where you want Amazon Macie to archive findings automatically. Suppression rules are helpful in situations where you've reviewed a class of findings and don't want to be notified of them again. When you create a suppression rule, you specify filter criteria, a name, and, optionally, a description of the rule. Macie then uses the rule's criteria to determine which findings to archive automatically. By using suppression rules, you can streamline your analysis of findings.

If you suppress findings with a suppression rule, Macie continues to generate findings for subsequent occurrences of sensitive data and potential policy violations that match the rule's criteria. However, Macie automatically changes the status of the findings to *archived*. This means that the findings don't appear by default on the Amazon Macie console, but they persist in Macie until they expire. (Macie stores findings for 90 days.) This also means that Macie doesn't publish the findings to Amazon EventBridge as events or to AWS Security Hub.

Note that suppression rules might work differently for your account, if your account is part of an organization that centrally manages multiple Macie accounts. This depends on the category of findings that you want to suppress, and whether you have a Macie administrator or member account:

- **Policy findings** – Only a Macie administrator can suppress policy findings for the organization's accounts.

  If you have a Macie administrator account and you create a suppression rule, Macie applies the rule to policy findings for all the accounts in your organization unless you configure the rule to exclude specific accounts. If you have a member account and you want to suppress policy findings for your account, work with your Macie administrator to suppress the findings.
- **Sensitive data findings** – A Macie administrator and individual members can suppress sensitive data findings that their sensitive data discovery jobs produce. A Macie administrator can also suppress findings that Macie generates while performing automated sensitive data discovery for the organization.

  Only the account that creates a sensitive data discovery job can suppress or otherwise access sensitive data findings that the job produces. Only the Macie administrator account for an organization can suppress or otherwise access findings that automated sensitive data discovery produces for accounts in the organization.

For more information about the tasks that administrators and members can perform, see Macie administrator and member account relationships.

Also note that suppression rules are different from filter rules. A *filter rule* is a set of filter criteria that you create and save to use again when you review findings on the Amazon Macie console. Although both types of rules store and apply filter criteria, a filter rule doesn't perform any action on findings that match the rule's criteria. Instead, a filter rule only determines which findings appear on the console after you apply the rule. For more information, see Defining filter rules. Depending on your analysis goals, you might determine that it's best to create a filter rule instead of a suppression rule.

To create a suppression rule for findings

You can create a suppression rule by using the Amazon Macie console or the Amazon Macie API. Before you create a suppression rule, it's important to note that you can't restore (unarchive) findings that you suppress using a suppression rule. You can, however, review suppressed findings by using Macie.

Console

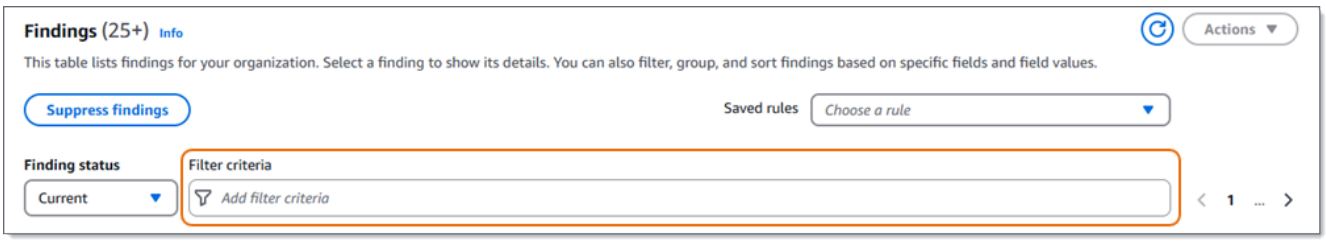
Follow these steps to create a suppression rule by using the Amazon Macie console.

To create a suppression rule

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Findings**.

> **Tip**
>
> To use an existing suppression or filter rule as a starting point, choose the rule from the **Saved rules** list.
> You can also streamline creation of a rule by first pivoting and drilling down on findings by a predefined logical group. If you do this, Macie automatically creates and applies the appropriate filter conditions, which can be a helpful starting point for creating a rule. To do this, choose **By bucket**, **By type**, or **By job** in the navigation pane (under **Findings**). Then choose an item in the table. In the details panel, choose the link for the field to pivot on.

3. In the **Filter criteria** box, add filter conditions that specify attributes of the findings that you want the rule to suppress.



To learn how to add filter conditions, see Creating and applying filters to Macie findings.

4. When you finish adding filter conditions for the rule, choose **Suppress findings**.
5. Under **Suppression rule**, enter a name and, optionally, a description of the rule.
6. Choose **Save**.

API

To create a suppression rule programmatically, use the CreateFindingsFilter operation of the Amazon Macie API and specify the appropriate values for the required parameters:

- For the `action` parameter, specify `ARCHIVE` to ensure that Macie suppresses findings that match the criteria of the rule.
- For the `criterion` parameter, specify a map of conditions that define the filter criteria for the rule.

  In the map, each condition should specify a field, an operator, and one or more values for the field. The type and number of values depends on the field and operator that you choose. For information about the fields, operators, and types of values that you can use in a condition, see: Fields for filtering Macie findings, Using operators in conditions, and Specifying values for fields.

To create a suppression rule by using the AWS Command Line Interface (AWS CLI), run the create-findings-filter command and specify the appropriate values for the required parameters. The following examples create a suppression rule that returns all sensitive data findings that are in the current AWS Region and report occurrences of mailing addresses (and no other types of sensitive data) in S3 objects.

This example is formatted for Linux, macOS, or Unix, and it uses the backslash (\) line-continuation character to improve readability.

```
$ aws macie2 create-findings-filter \
```

```
--action ARCHIVE \
--name my_suppression_rule \
--finding-criteria '{"criterion":
{"classificationDetails.result.sensitiveData.detections.type":{"eqExactMatch":
["ADDRESS"]}}}'
```

This example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws macie2 create-findings-filter ^
--action ARCHIVE ^
--name my_suppression_rule ^
--finding-criteria={\"criterion\":
{\"classificationDetails.result.sensitiveData.detections.type\":{\"eqExactMatch\":
[\"ADDRESS\"]}}}
```

Where:

- *my_suppression_rule* is the custom name for the rule.
- `criterion` is a map of filter conditions for the rule:
  - *classificationDetails.result.sensitiveData.detections.type* is the JSON name of the **Sensitive data detection type** field.
  - *eqExactMatch* specifies the *equals exact match* operator.
  - *ADDRESS* is an enumerated value for the **Sensitive data detection type** field.

If the command runs successfully, you receive output similar to the following.

```
{
    "arn": "arn:aws:macie2:us-west-2:123456789012:findings-filter/8a3c5608-
aa2f-4940-b347-d1451example",
    "id": "8a3c5608-aa2f-4940-b347-d1451example"
}
```

Where `arn` is the Amazon Resource Name (ARN) of the suppression rule that was created, and `id` is the unique identifier for the rule.

For additional examples of filter criteria, see Filtering findings programmatically with the Amazon Macie API.

Reviewing suppressed findings in Macie

If you suppress findings with a suppression rule, Amazon Macie continues to generate findings for subsequent occurrences of sensitive data and potential policy violations that match the rule's criteria. However, Macie automatically changes the status of the findings to *archived*. This means that the findings don't appear by default on the Amazon Macie console, but they persist in Macie until they expire. (Macie stores findings for 90 days.) This also means that Macie doesn't publish the findings to Amazon EventBridge as events or to AWS Security Hub.

Because suppressed findings persist in Macie for up to 90 days, you can access and review them before they expire. In addition to broadening your analysis of findings, this can help you determine whether to adjust your suppression criteria. To adjust the criteria, change the suppression rules for your account.

You can review suppressed findings on the Amazon Macie console by changing your filter settings.

To review suppressed findings on the console

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Findings**. The **Findings** page displays findings that Macie created or updated for your account in the current AWS Region during the past 90 days. By default, this doesn't include findings that were suppressed by a suppression rule.
3. To pivot on and review the findings by a predefined logical group, choose **By bucket**, **By type**, or **By job** in the navigation pane (under **Findings**).
4. For **Finding status**, do one of the following:

   - To display only suppressed findings, choose **Archived**.
   - To display both suppressed and unsuppressed findings, choose **All**.
   - To hide suppressed findings again, choose **Current**.

You can also access suppressed findings by using the Amazon Macie API. To retrieve a list of suppressed findings, use the ListFindings operation. In your request, include a filter condition that specifies `true` for the `archived` field. For an example of how to do this by using the AWS Command Line Interface (AWS CLI), see Filtering findings programmatically. To then retrieve the details of one or more suppressed findings, use the GetFindings operation. In your request, specify the unique identifier for each finding to retrieve.

Note

As you review the findings, note that suppression rules can work differently for accounts that are part of an organization. This depends on a finding's category and whether you have a Macie administrator or member account:

- **Policy findings** – Only a Macie administrator can suppress policy findings for the organization's accounts.

  If you have a Macie administrator account and you created a suppression rule, Macie applies the rule to policy findings for all the accounts in your organization unless you configured the rule to exclude specific accounts. If you have a member account and you want to suppress policy findings for your account, work with your Macie administrator to suppress the findings.
- **Sensitive data findings** – A Macie administrator and individual members can suppress sensitive data findings that their sensitive data discovery jobs produce. A Macie administrator can also suppress findings that Macie generates while performing automated sensitive data discovery for the organization.

  Only the account that creates a sensitive data discovery job can suppress or otherwise access sensitive data findings that the job produces. Only the Macie administrator account for an organization can suppress or otherwise access findings that automated sensitive data discovery produces for accounts in the organization.

For more information about the tasks that administrators and members can perform, see Macie administrator and member account relationships.

Changing a suppression rule for Macie findings

After you create a suppression rule, you can change the settings for the rule. A *suppression rule* is a set of attribute-based filter criteria that defines cases where you want Amazon Macie to archive findings automatically. Suppression rules are helpful in situations where you've reviewed a class of findings and don't want to be notified of them again. Each rule consists of a set of filter criteria, a name, and, optionally, a description.

If you change the criteria of a suppression rule, findings that were previously suppressed by the rule continue to be suppressed. The findings continue to have a status of *archived* and Macie

doesn't publish them to Amazon EventBridge or AWS Security Hub. Macie applies the new criteria only to new sensitive data findings, new policy findings, and subsequent occurrences of existing policy findings.

In addition to changing the criteria or other settings for a rule, you can assign tags to a rule. A *tag* is a label that you define and assign to certain types of AWS resources. Each tag consists of a required tag key and an optional tag value. Tags can help you identify, categorize, and manage resources in different ways, such as by purpose, owner, environment, or other criteria. To learn more, see Tagging Macie resources.

To change a suppression rule for findings

To assign tags or change the settings for a suppression rule, you can use the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to assign tags or change the settings for a suppression rule by using the Amazon Macie console.

To change a suppression rule

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Findings**.
3. In the **Saved rules** list, choose the edit icon () next to the suppression rule that you want to change or assign tags to.
4. Do any of the following:
   - To change the criteria of the rule, use the **Filter criteria** box. In the box, enter conditions that specify attributes of the findings that you want the rule to suppress. To learn how, see Creating and applying filters to Macie findings.
   - To change the name of the rule, enter a new name in the **Name** box under **Suppression rule**.
   - To change the description of the rule, enter a new description in the **Description** box under **Suppression rule**.
   - To assign tags to the rule, choose **Manage tags** under **Suppression rule**. Then add, review, and change the tags as necessary. A rule can have as many as 50 tags.

5. When you finish making changes, choose **Save**.

API

To change a suppression rule programmatically, use the UpdateFindingsFilter operation of the Amazon Macie API. When you submit your request, use the supported parameters to specify a new value for each setting that you want to change.

For the `id` parameter, specify the unique identifier for the rule to change. You can get this identifier by using the ListFindingsFilter operation to retrieve a list of suppression and filter rules for your account. If you're using the AWS Command Line Interface (AWS CLI), run the list-findings-filters command to retrieve this list.

To change a suppression rule by using the AWS CLI, run the update-findings-filter command and use the supported parameters to specify a new value for each setting that you want to change. For example, the following command changes the name of an existing suppression rule.

```
C:\> aws macie2 update-findings-filter --id 8a3c5608-aa2f-4940-b347-d1451example --
name mailing_addresses_only
```

Where:

- *8a3c5608-aa2f-4940-b347-d1451example* is the unique identifier for the rule.
- *mailing_addresses_only* is the new name for the rule.

If the command runs successfully, you receive output similar to the following.

```
{
    "arn": "arn:aws:macie2:us-west-2:123456789012:findings-filter/8a3c5608-
aa2f-4940-b347-d1451example",
    "id": "8a3c5608-aa2f-4940-b347-d1451example"
}
```

Where `arn` is the Amazon Resource Name (ARN) of the rule that was changed, and `id` is the unique identifier for the rule.

Similarly, the following example converts a filter rule to a suppression rule by changing the value for the `action` parameter from `NOOP` to `ARCHIVE`.

```
C:\> aws macie2 update-findings-filter --id 8a1c3508-aa2f-4940-b347-d1451example --
action ARCHIVE
```

Where:

- *8a1c3508-aa2f-4940-b347-d1451example* is the unique identifier for the rule.
- *ARCHIVE* is the new action for Macie to perform on findings that match the criteria of the rule—suppress the findings.

If the command runs successfully, you receive output similar to the following:

```
{
    "arn": "arn:aws:macie2:us-west-2:123456789012:findings-filter/8a1c3508-
aa2f-4940-b347-d1451example",
    "id": "8a1c3508-aa2f-4940-b347-d1451example"
}
```

Where `arn` is the Amazon Resource Name (ARN) of the rule that was changed, and `id` is the unique identifier for the rule.

Deleting a suppression rule for Macie findings

You can delete a suppression rule at any time. If you delete a suppression rule, Amazon Macie stops suppressing new and subsequent occurrences of findings that match the rule's criteria and aren't suppressed by other rules. Note, however, that Macie might continue to suppress findings that it's currently processing and match the rule's criteria.

After you delete a suppression rule, new and subsequent occurrences of findings that match the rule's criteria have a status of *current* (not *archived*). This means that they appear by default on the Amazon Macie console. In addition, Macie publishes them to Amazon EventBridge as events. Depending on the publication settings for your account, Macie also publishes the findings to AWS Security Hub.

To delete a suppression rule for findings

You can delete a suppression rule by using the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to delete a suppression rule by using the Amazon Macie console.

To delete a suppression rule

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Findings**.
3. In the **Saved rules** list, choose the edit icon () next to the suppression rule that you want to delete.
4. Under **Suppression rule**, choose **Delete**.

API

To delete a suppression rule programmatically, use the DeleteFindingsFilter operation of the Amazon Macie API. For the `id` parameter, specify the unique identifier for the suppression rule to delete. You can get this identifier by using the ListFindingsFilter operation to retrieve a list of suppression and filter rules for your account. If you're using the AWS Command Line Interface (AWS CLI), run the list-findings-filters command to retrieve this list.

To delete a suppression rule by using the AWS CLI, run the delete-findings-filter command. For example:

```
C:\> aws macie2 delete-findings-filter --id 8a3c5608-aa2f-4940-b347-d1451example
```

Where *8a3c5608-aa2f-4940-b347-d1451example* is the unique identifier for the suppression rule to delete.

If the command runs successfully, Macie returns an empty HTTP 200 response. Otherwise, Macie returns an HTTP 4*xx* or 500 response that indicates why the operation failed.

Monitoring and processing Macie findings

To support integration with other applications, services, and systems, such as monitoring or event management systems, Amazon Macie automatically publishes policy and sensitive data findings to Amazon EventBridge as events. For additional support and broader analysis of your organization's security posture, you can configure Macie to also publish policy and sensitive data findings to AWS Security Hub.

Amazon EventBridge

Amazon EventBridge, formerly Amazon CloudWatch Events, is a serverless event bus service that delivers a stream of real-time data from applications and services, and routes that data to targets such as AWS Lambda functions, Amazon Simple Notification Service topics, and Amazon Kinesis streams. With EventBridge, you can automate monitoring and processing of certain types of events, including events that Macie publishes for findings. To learn more, see Processing findings with Amazon EventBridge.

If you integrate AWS User Notifications with Macie, you can also use EventBridge events to automatically generate notifications about events that Macie publishes for findings. With User Notifications, you create custom rules and configure delivery channels for receiving notifications about EventBridge events of interest. The delivery channels include email, Amazon Q Developer in chat applications, and push notifications in the AWS Console Mobile Application. You can also review notifications in a central location on the AWS Management Console. To learn more, see Monitoring findings with AWS User Notifications.

AWS Security Hub

AWS Security Hub is a security service that provides you with a comprehensive view of your security state across your AWS environment. It collects security data from AWS services and supported AWS Partner Network security solutions, and helps you check your environment against security industry standards and best practices. It also helps you analyze security trends and identify high-priority issues.

With Security Hub, you can review and evaluate Macie findings as part of a broader analysis of your organization's security posture. You can also aggregate findings from multiple AWS Regions, and monitor and process aggregated findings data from a single Region. To learn more, see Evaluating findings with AWS Security Hub.

When Macie creates a finding, it automatically publishes the finding to EventBridge as a new event. Depending on the publication settings that you choose for your account, Macie can also publish the finding to Security Hub. Macie publishes each new finding immediately after it finishes processing the finding. If Macie detects a subsequent occurrence of an existing policy finding, it publishes an update to the existing EventBridge event for the finding. Depending on your publication settings, Macie can also publish the update to Security Hub. Macie publishes these updates on a recurring basis, using a publication frequency that you specify in the publication settings for your account.

In addition to the preceding options, you can query and retrieve findings data directly by using the Amazon Macie API. The Amazon Macie API gives you comprehensive, programmatic access to the data. To query the data, you can send HTTPS requests directly to Macie or use a current version of an AWS SDK or an AWS command line tool. If you query the data, Macie returns the results in a JSON response. You can then pass the results to another service or application for additional processing, monitoring, or reporting. For more information, see the Amazon Macie API Reference.

Topics

Configuring publication settings for Macie findings

To support integration with other applications, services, and systems, Amazon Macie automatically publishes both policy findings and sensitive data findings to Amazon EventBridge as events. For information about how you can use EventBridge to monitor and process findings, see Processing findings with Amazon EventBridge.

You can configure Macie to automatically publish findings to AWS Security Hub too, using destination options that you specify in the publication settings for your account. With these options, you can configure Macie to publish only policy findings, only sensitive data findings, or both policy and sensitive data findings to Security Hub. You can also configure Macie to stop publishing any findings to Security Hub. For information about how you can use Security Hub to evaluate and process findings, see Evaluating findings with AWS Security Hub.

For policy findings, the timing with which Macie publishes a finding to another AWS service depends on whether the finding is new and the publication frequency that you specify for your account. For sensitive data findings, the timing is always immediate—Macie publishes a sensitive data finding immediately after it finishes processing the finding. Unlike policy findings, Macie treats all sensitive data findings as new (unique).

Note that Macie doesn't publish policy or sensitive data findings that are archived automatically by a suppression rule. In other words, Macie doesn't publish suppressed findings to other AWS services.

Topics

Choosing publication destinations for findings

You can configure Amazon Macie to automatically publish policy and sensitive data findings to AWS Security Hub in addition to Amazon EventBridge. By default, Macie publishes only new and updated policy findings to Security Hub. To change or extend the default configuration, adjust the publication destination settings for your account.

When you adjust your destination settings, you choose the categories of findings that you want Macie to publish to Security Hub—only policy findings, only sensitive data findings, or both policy and sensitive data findings. You can also choose to stop publishing any category of finding to Security Hub.

If you change your destination settings, your change applies only to the current AWS Region. If you're the Macie administrator for an organization, your change applies only to your account. It doesn't apply to any member accounts in your organization. For more information, see Managing multiple accounts.

To choose publication destinations for findings

Follow these steps to change your destination settings by using the Amazon Macie console. To do this programmatically, use the PutFindingsPublicationConfiguration operation of the Amazon Macie API.

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.

2. In the navigation pane, choose **Settings**.
3. In the **Publication of findings** section, under **Destinations**, choose from the following options:

   - **Publish policy findings to Security Hub** – Select this checkbox to start publishing new and updated policy findings to Security Hub automatically. To stop publishing new and updated policy findings to Security Hub, clear this checkbox.

     If you select this checkbox and you have existing policy findings, Macie doesn't publish them to Security Hub. Instead, Macie publishes only those policy findings that it creates or updates after you save your change.
   - **Publish sensitive data findings to Security Hub** – Select this checkbox to start publishing new sensitive data findings to Security Hub automatically. To stop publishing new sensitive data findings to Security Hub, clear this checkbox.

     If you select this checkbox and you have existing sensitive data findings, Macie doesn't publish them to Security Hub. Instead, Macie publishes only those sensitive data findings that it creates after you save your change.
4. Choose **Save**.

If you chose to publish any category of finding to Security Hub, make sure that you also enable Security Hub in the current Region and configure it to accept findings from Macie. Otherwise, you won't be able to access the findings in Security Hub. To learn how to accept findings in Security Hub, see Enabling and managing integrations in the *AWS Security Hub User Guide*.

Changing the publication frequency for findings

In Amazon Macie, each finding has a unique identifier. Macie uses this identifier to determine when to publish a finding to another AWS service:

- **New findings** – When Macie creates a new policy or sensitive data finding, it assigns a unique identifier to the finding as part of processing the finding. Immediately after Macie finishes processing the finding, it publishes the finding to Amazon EventBridge as a new event. Depending on the publication settings for your account, Macie also publishes the finding as a new finding in AWS Security Hub.
- **Updated findings** – When Macie detects a subsequent occurrence of an existing policy finding, it updates the existing finding by adding details about the subsequent occurrence and incrementing the count of occurrences. Macie also publishes these updates to the existing

EventBridge event and, depending on the publication settings for your account, the existing Security Hub finding. By default, Macie publishes updates every 15 minutes as part of a recurring publication cycle. This means any policy findings that are updated after the most recent publication cycle will be held, updated again as necessary, and included in the next publication cycle (approximately 15 minutes later).

You can change the frequency with which Macie publishes updates to existing policy findings in other AWS services. For example, you might configure Macie to publish the updates every hour. If you do this and a publication occurs at 12:00, any updates that occur after 12:00 are published at 13:00.

If you change the frequency, your change applies only to the current AWS Region. If you're the Macie administrator for an organization, your change also applies to all member accounts in your organization. For more information, see Managing multiple accounts.

To change the publication frequency for updated findings

Follow these steps to change the publication frequency by using the Amazon Macie console. To do this programmatically, use the UpdateMacieSession operation of the Amazon Macie API.

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Settings**.
3. In the **Publication of findings** section, under **Update frequency for policy findings**, choose how often you want Macie to publish updates to policy findings in other AWS services.
4. Choose **Save**.

Processing Macie findings with Amazon EventBridge

Amazon EventBridge, formerly Amazon CloudWatch Events, is a serverless event bus service. EventBridge delivers a stream of real-time data from applications and services, and routes that data to targets such as AWS Lambda functions, Amazon Simple Notification Service (Amazon SNS) topics, and Amazon Kinesis streams. To learn more about EventBridge, see the Amazon EventBridge User Guide.

With EventBridge, you can automate monitoring and processing of certain types of events. This includes events that Amazon Macie publishes automatically for new policy findings and sensitive data findings. This also includes events that Macie publishes automatically for subsequent

occurrences of existing policy findings. For details about how and when Macie publishes these events, see Configuring publication settings for findings.

By using EventBridge and the events that Macie publishes for findings, you can monitor and process findings in near real time. You can then act upon findings by using other applications and services. For example, you might use EventBridge to send specific types of new findings to an AWS Lambda function. The Lambda function might then process and send the data to your security incident and event management (SIEM) system. If you integrate AWS User Notifications with Macie, you can also use the events to be notified of findings automatically through delivery channels that you specify.

In addition to automated monitoring and processing, use of EventBridge enables longer-term retention of your findings data. Macie stores findings for 90 days. With EventBridge, you can send findings data to your preferred storage platform and store the data for as long as you like.

> **Note**
>
> For long-term retention, also configure Macie to store your sensitive data discovery results in an S3 bucket. A *sensitive data discovery result* is a record that logs details about the analysis that Macie performed on an S3 object to determine whether the object contains sensitive data. To learn more, see Storing and retaining sensitive data discovery results.

Topics

Working with Amazon EventBridge

With Amazon EventBridge, you create rules to specify which events you want to monitor and which targets you want to perform automated actions for those events. A *target* is a destination that EventBridge sends events to.

To automate monitoring and processing tasks for findings, you can create an EventBridge rule that automatically detects Amazon Macie finding events and sends those events to another application or service for processing or other action. You can tailor the rule to send only those events that meet certain criteria. To do this, specify criteria that derive from the Amazon EventBridge event schema for Macie findings.

For example, you can create a rule that sends specific types of new findings to an AWS Lambda function. The Lambda function can then perform tasks such as: process and send the data to your SIEM system; automatically apply a certain type of server-side encryption to an S3 object; or, restrict access to an S3 object by changing the object's access control list (ACL). Or you can create a rule that automatically sends new high-severity findings to an Amazon SNS topic, which then notifies your incident response team of the finding.

In addition to invoking Lambda functions and notifying Amazon SNS topics, EventBridge supports other types of targets and actions, such as relaying events to Amazon Kinesis streams, activating AWS Step Functions state machines, and invoking the AWS Systems Manager run command. For information about supported targets, see Event bus targets in the *Amazon EventBridge User Guide*.

Creating Amazon EventBridge rules for Macie findings

The following procedures explain how to use the Amazon EventBridge console and the AWS Command Line Interface (AWS CLI) to create an EventBridge rule for Amazon Macie findings. The rule detects EventBridge events that use the event schema and pattern for Macie findings, and it sends those events to an AWS Lambda function for processing.

AWS Lambda is a compute service that you can use to run code without provisioning or managing servers. You package your code and upload it to AWS Lambda as a *Lambda function*. AWS Lambda then runs the function when the function is invoked. A function can be invoked manually by you, automatically in response to events, or in response to requests from applications or services. For information about creating and invoking Lambda functions, see the AWS Lambda Developer Guide.

Console

Follow these steps to use the Amazon EventBridge console to create a rule that automatically sends all Macie finding events to a Lambda function for processing. The rule uses default settings for rules that run when specific events are received. For details about rule settings or to learn how to create a rule that uses custom settings, see Creating rules that react to events in the *Amazon EventBridge User Guide*.

> **Tip**
>
> You can also create a rule that uses a custom pattern to detect and act upon only a subset of Macie finding events. This subset can be based on specific fields that Macie includes in a finding event. To learn about the available fields, see Amazon EventBridge

event schema for Macie findings. To learn about using custom patterns in rules, see Creating event patterns in the *Amazon EventBridge User Guide*.

Before you create this rule, create the Lambda function that you want the rule to use as a target. When you create the rule, you'll need to specify this function as the target for the rule.

To create an event rule by using the console

1. Open the Amazon EventBridge console at https://console.aws.amazon.com/events/.
2. In the navigation pane, under **Buses**, choose **Rules**.
3. In the **Rules** section, choose **Create rule**.
4. On the **Define rule detail** page, do the following:
   - For **Name**, enter a name for the rule.
   - (Optional) For **Description**, enter a brief description of the rule.
   - For **Event bus**, ensure that **default** is selected and **Enable the rule on the selected event bus** is turned on.
   - For **Rule type**, choose **Rule with an event pattern**.
5. When you finish, choose **Next**.
6. On the **Build event pattern** page, do the following:
   - For **Event source**, choose **AWS events or EventBridge partner events**.
   - (Optional) For **Sample event**, review a sample finding event for Macie to learn what an event might contain. To do this, choose **AWS events**. Then, for **Sample events**, choose **Macie Finding**.
   - For **Creation method**, choose **Use pattern form**.
   - For **Event pattern**, enter the following settings:
     - For **Event source**, choose **AWS services**.
     - For **AWS service**, choose **Macie**.
     - For **Event type**, choose **Macie Finding**.
7. When you finish, choose **Next**.
8. On the **Select targets** page, do the following:
   - For **Target types**, choose **AWS service**.

- For **Select a target**, choose **Lambda function**. Then, for **Function**, choose the Lambda function that you want to send finding events to.
- For **Configure version/alias**, enter version and alias settings for the target Lambda function.
- (Optional) For **Additional settings**, enter custom settings to specify which event data you want to send to the Lambda function. You can also specify how to handle events that aren't delivered to the function successfully.

9. When you finish, choose **Next**.
10. On the **Configure tags** page, optionally enter one or more tags to assign to the rule. Then choose **Next**.
11. On the **Review and create** page, review the rule's settings and verify that they're correct.

    To change a setting, choose **Edit** in the section that contains the setting, and then enter the correct setting. You can also use the navigation tabs to go to the page that contains a setting.
12. When you finish verifying the settings, choose **Create rule**.

AWS CLI

Follow these steps to use the AWS CLI to create an EventBridge rule that sends all Macie finding events to a Lambda function for processing. The rule uses default settings for rules that run when specific events are received. In this procedure, the commands are formatted for Microsoft Windows. For Linux, macOS, or Unix, replace the caret (^) line-continuation character with a backslash (\).

Before you create this rule, create the Lambda function that you want the rule to use as a target. When you create the function, note the Amazon Resource Name (ARN) of the function. You'll need to enter this ARN when you specify the target for the rule.

To create an event rule by using the AWS CLI

1. Create a rule that detects events for all the findings that Macie publishes to EventBridge. To do this, run the EventBridge put-rule command. For example:

```
C:\> aws events put-rule ^
--name MacieFindings ^
--event-pattern "{\"source\":[\"aws.macie\"]}"
```

Where *MacieFindings* is the name that you want for the rule.

> **Tip**
>
> You can also create a rule that uses a custom pattern (`event-pattern`) to detect and act upon only a subset of Macie finding events. This subset can be based on specific fields that Macie includes in a finding event. To learn about the available fields, see Amazon EventBridge event schema for Macie findings. To learn about using custom patterns in rules, see Creating event patterns in the *Amazon EventBridge User Guide*.

If the command runs successfully, EventBridge responds with the ARN of the rule. Note this ARN. You'll need to enter it in step 3.

2. Specify the Lambda function to use as a target for the rule. To do this, run the EventBridge put-targets command. For example:

```
C:\> aws events put-targets ^
--rule MacieFindings ^
--targets Id=1,Arn=arn:aws:lambda:regionalEndpoint:accountID:function:my-findings-function
```

Where *MacieFindings* is the name that you specified for the rule in step 1, and the value for the `Arn` parameter is the ARN of the function that you want the rule to use as a target.

3. Add permissions that allow the rule to invoke the target Lambda function. To do this, run the Lambda add-permission command. For example:

```
C:\> aws lambda add-permission ^
--function-name my-findings-function ^
--statement-id Sid ^
--action lambda:InvokeFunction ^
--principal events.amazonaws.com ^
--source-arn arn:aws:events:regionalEndpoint:accountId:rule:MacieFindings
```

Where:

- *my-findings-function* is the name of the Lambda function that you want the rule to use as a target.
- *Sid* is a statement identifier that you define to describe the statement in the Lambda function policy.
- `source-arn` is the ARN of the EventBridge rule.

If the command runs successfully, you receive output similar to the following:

```
{
  "Statement": "{\"Sid\":\"sid\",
    \"Effect\":\"Allow\",
    \"Principal\":{\"Service\":\"events.amazonaws.com\"},
    \"Action\":\"lambda:InvokeFunction\",
    \"Resource\":\"arn:aws:lambda:us-east-1:111122223333:function:my-findings-
function\",
    \"Condition\":
      {\"ArnLike\":
        {\"AWS:SourceArn\":
         \"arn:aws:events:us-east-1:111122223333:rule/MacieFindings\"}}}"
}
```

The `Statement` value is a JSON string version of the statement that was added to the Lambda function policy.

Monitoring Macie findings with AWS User Notifications

AWS User Notifications is a service that acts as a central location for your AWS notifications on the AWS Management Console. This includes notifications such as Amazon CloudWatch alarms, AWS Support cases, and communications from other AWS services. With User Notifications, you can configure custom rules and delivery channels for receiving notifications about certain types of Amazon EventBridge events. The delivery channels include email, Amazon Q Developer in chat applications, and push notifications in the AWS Console Mobile Application. You can also review notifications on the AWS User Notifications console. To learn more about User Notifications, see the AWS User Notifications User Guide.

Amazon Macie integrates with AWS User Notifications, which means you can configure User Notifications to notify you of events that Macie publishes to EventBridge for policy and sensitive

data findings. If a finding event matches criteria that you specify, User Notifications generates a notification. The notification includes key details of the associated finding, such as the finding's type and severity, and the name of the affected resource. User Notifications can also send the notification to one or more delivery channels that you specify. You can tailor your choice of delivery channels to align with your security and compliance workflows.

For example, you might configure User Notifications to generate notifications for specific types of new, high-severity findings. You might also specify Amazon Q Developer in chat applications as a delivery channel for those notifications. User Notifications then detects EventBridge events for the findings, generates notifications that include data from the findings, and sends the notifications to Amazon Q Developer in chat applications. Amazon Q Developer in chat applications might then route the notifications to a Slack channel or an Amazon Chime chat room to notify your incident response team.

Topics

Working with AWS User Notifications

With AWS User Notifications, you create rules to specify the types of Amazon EventBridge events that you want to monitor and receive notifications for. A rule defines criteria that an EventBridge event must match in order to generate a notification. You can also choose one or more delivery channels for a rule. Delivery channels specify where you want to receive notifications for events that match a rule's criteria.

If User Notifications detects an EventBridge event that matches a rule's criteria, it performs the following general tasks:

1. Extracts a subset of data from the event.
2. Generates a notification that contains the extracted data.
3. Sends the notification to delivery channels that you specify for that type of event.

The design and structure of the notification is optimized for each delivery channel that it's sent to.

To control the frequency or number of notifications that you receive, you can configure aggregation settings for a rule. If you enable these settings, User Notifications combines data for multiple events into a single notification. You can choose to send aggregated event notifications quickly and frequently, which you might want to do for high-severity finding events. Or send them less frequently to receive fewer notifications, which you might want to do for low-severity finding events. If you combine event data, you can drill down to review the details of each aggregated event by using the AWS User Notifications console. From there, you can also navigate to each associated finding on the Amazon Macie console.

Enabling and configuring AWS User Notifications for Macie findings

To enable AWS User Notifications to generate notifications for Amazon Macie findings, create a notification configuration for Macie in User Notifications. A *notification configuration* specifies the criteria for a rule. It also specifies delivery channels and other settings for monitoring and sending notifications about Amazon EventBridge events that match the rule's criteria. For detailed information about creating a notification configuration, see Getting started with AWS User Notifications in the *AWS User Notifications User Guide*.

To create a notification configuration for Macie findings, choose the following options for the event rule:

- For **AWS service name**, choose **Macie**.
- For **Event type**, choose **Macie Finding**.
- For **Regions**, select each AWS Region in which you use Macie and want to be notified of findings.

With this configuration, User Notifications monitors EventBridge events for your AWS account and generates notifications for all Macie finding events in the Regions that you selected. The events match the following criteria:

- `source` equals `aws.macie`
- `detail-type` equals `Macie Finding`

The underlying JSON pattern for the event rule is:

```
{
```

```
    "source": ["aws.macie"],
    "detail-type": ["Macie Finding"]
}
```

To refine the rule and generate notifications only for a subset of findings, you can customize the JSON pattern for the rule. To do this, specify additional criteria that derive from the Amazon EventBridge event schema for Macie findings.

If you create a rule that uses a custom JSON pattern, you can create multiple notification configurations for Macie findings. You can then tailor the delivery channels and other settings for each configuration to align with your security and compliance workflows for specific types of findings.

For example, you might create one rule that notifies you if Macie generates or updates a *Policy:IAMUser/S3BucketPublic* finding. In this case, the pattern for the rule might be:

```
{
    "source": ["aws.macie"],
    "detail-type": ["Macie Finding"],
    "detail": {
        "type": ["Policy:IAMUser/S3BucketPublic"]
    }
}
```

And you might create another rule that notifies you if Macie generates a sensitive data finding for an S3 bucket that's publicly accessible. In this case, the pattern for the rule might be:

```
{
    "source": ["aws.macie"],
    "detail-type": ["Macie Finding"],
    "detail": {
        "type": [ { "prefix": "SensitiveData" } ],
        "resourcesAffected": {
            "effectivePermission": ["PUBLIC"]
        }
    }
}
```

If you create multiple notification configurations for Macie findings, it's a good idea to ensure that the rule for each configuration is unique. Otherwise, you might receive duplicate notifications for individual findings.

To learn more about customizing event patterns for rules, see Using customized JSON event patterns in the *AWS User Notifications User Guide*.

Mapping AWS User Notifications fields to Macie finding fields

When AWS User Notifications generates a notification for an Amazon Macie finding, it populates the notification with data from a subset of fields in the corresponding Amazon EventBridge event. These fields provide key details of the associated finding, such as the finding's type and severity, and the name of the affected resource.

If you review a notification on the AWS User Notifications console, the notification includes all the data for this subset of fields. It also provides a link to the associated finding on the Amazon Macie console. If you review a notification in other delivery channels, it might contain data for only some of the fields. This is because User Notifications tailors the design and structure of its notifications to work with each type of delivery channel that it supports.

The following table lists the fields that might be included in a notification for a finding. In the table, the **Notification field** column describes (in *italics*) or indicates the name of a field in a notification. The **Finding event field** column uses dot notation to indicate the name of the corresponding JSON field in an EventBridge event for a finding. The **Description** column describes the data that's stored in the field.

Notification field	Finding event field	Description
Message headline	`detail.type`	The finding's type. For example: `Policy:IAMUser/S3BucketPublic` or `Sensitive Data:S3Object/Financial` .
Summary	`detail.title`	The brief description of the finding. For example: `The S3 object contains financial information.`

Notification field	Finding event field	Description
Description	`detail.description`	The full description of the finding. For example: `The S3 object contains financial information such as bank account numbers or credit card numbers.`
Severity	`detail.severity.description`	The qualitative representation of the finding's severity: `Low`, `Medium`, or `High`.
Finding ID	`detail.id`	The unique identifier for the finding.
Created	`detail.createdAt`	The date and time when Macie created the finding.
Updated	`detail.updatedAt`	The date and time when Macie most recently updated the finding. For sensitive data findings, this value is the same as the value for the *Created* (`detail.createdAt`) field. All sensitive data findings are considered new (unique).
Affected S3 bucket	`detail.resourcesAffected.s3Bucket.arn`	The Amazon Resource Name (ARN) of the affected S3 bucket.

Notification field	Finding event field	Description
Affected S3 object	`detail.resourcesAffected.s3Object.path`	The name (*key*) of the affected S3 object, including the name of the bucket that stores the object and, if applicable, the object's prefix. This field isn't included in notifications for policy findings.

Notification field	Finding event field	Description

Sensitive data detections	`detail.classificationDetails.result.sensitiveData.detections...`
	
	And/Or
	
	`detail.classificationDetails.result.customDataIdentifiers.detections...`
		This is a concatenation of multiple fields in an event for a sensitive data finding. This field isn't included in notifications for policy findings.
		
		If a managed data identifier detected the sensitive data, this field specifies the category, type, and number (`count`) of occurrences of the sensitive data that was detected. For example: `PERSONAL_INFORMATION: USA_SOCIAL_SECURITY_NUMBER 100 occurrences .`
		
		If a custom data identifier detected the sensitive data, this field specifies the name of the custom data identifier and the number (`count`) of occurrences of the sensitive data that was detected. For example: `Employee ID 20 occurrences .`
		
		If a finding reports multiple types of sensitive data, the notification includes data for up to four types. The data is populated first by any applicable custom data identifiers and then by any

Notification field	Finding event field	Description
		applicable managed data identifiers.

Changing AWS User Notifications settings for Macie findings

You can change your AWS User Notifications settings for Amazon Macie findings at any time. To do this, edit the notification configuration in User Notifications. To learn how, see Managing notification configurations in the *AWS User Notifications User Guide*.

If you have multiple notification configurations for Macie findings, changing the settings for one configuration doesn't affect the settings for your other configurations. You can edit all or only some of your configurations.

Disabling AWS User Notifications for Macie findings

To stop generating and receiving notifications from AWS User Notifications for Amazon Macie findings, delete the notification configuration in User Notifications. To learn how, see Managing notification configurations in the *AWS User Notifications User Guide*.

If you have multiple notification configurations for Macie findings, deletion of one configuration doesn't affect your other configurations. You can delete all or only some of your configurations.

Evaluating Macie findings with AWS Security Hub

AWS Security Hub is a service that provides you with a comprehensive view of your security posture across your AWS environment and helps you check your environment against security industry standards and best practices. It does this partly by consuming, aggregating, organizing, and prioritizing findings from multiple AWS services and supported AWS Partner Network security solutions. Security Hub helps you analyze your security trends and identify the highest priority security issues. With Security Hub, you can also aggregate findings from multiple AWS Regions, and then evaluate and process all the aggregated findings data from a single Region. To learn more about Security Hub, see the AWS Security Hub User Guide.

Amazon Macie integrates with Security Hub, which means that you can publish findings from Macie to Security Hub automatically. Security Hub can then include those findings in its analysis of your security posture. In addition, you can use Security Hub to evaluate and process policy

and sensitive data findings as part of a larger, aggregated set of findings data for your AWS environment. In other words, you can evaluate Macie findings while performing broader analyses of your organization's security posture, and remediate findings as necessary. Security Hub reduces the complexity of addressing large volumes of findings from multiple providers. In addition, it uses a standard format for all findings, including findings from Macie. Use of this format, the *AWS Security Finding Format (ASFF)*, eliminates the need for you to perform time-consuming data conversion efforts.

Topics

How Macie publishes findings to AWS Security Hub

In AWS Security Hub, security issues are tracked as findings. Some findings come from issues that are detected by AWS services, such as Amazon Macie, or by supported AWS Partner Network security solutions. Security Hub also has a set of rules that it uses to detect security issues and generate findings.

Security Hub provides tools to manage findings from all of these sources. You can review and filter lists of findings and review the details of individual findings. To learn how, see Reviewing finding history and finding details in the *AWS Security Hub User Guide*. You can also track the status of an investigation into a finding. To learn how, see Setting the workflow status of findings in the *AWS Security Hub User Guide*.

All findings in Security Hub use a standard JSON format called the *AWS Security Finding Format (ASFF)*. The ASFF includes details about the source of an issue, the affected resources, and the current status of a finding. For more information, see AWS Security Finding Format (ASFF) in the *AWS Security Hub User Guide*.

Types of findings that Macie publishes to Security Hub

Depending on the publication settings that you choose for your Macie account, Macie can publish all the findings that it creates to Security Hub, both sensitive data findings and policy findings. For information about these settings and how to change them, see Configuring publication settings for

findings. By default, Macie publishes only new and updated policy findings to Security Hub. Macie doesn't publish sensitive data findings to Security Hub.

Sensitive data findings

If you configure Macie to publish sensitive data findings to Security Hub, Macie automatically publishes each sensitive data finding that it creates for your account and it does so immediately after it finishes processing the finding. Macie does this for all sensitive data findings that aren't archived automatically by a suppression rule.

If you're the Macie administrator for an organization, publication is limited to findings from sensitive data discovery jobs that you ran and automated sensitive data discovery activities that Macie performed for your organization. Only the account that creates a job can publish sensitive data findings that the job produces. Only the Macie administrator account can publish sensitive data findings that automated sensitive data discovery produces for their organization.

When Macie publishes sensitive data findings to Security Hub, it uses the AWS Security Finding Format (ASFF), which is the standard format for all findings in Security Hub. In the ASFF, the `Types` field indicates a finding's type. This field uses a taxonomy that's slightly different from the finding type taxonomy in Macie.

The following table lists the ASFF finding type for each type of sensitive data finding that Macie can create.

Macie finding type	ASFF finding type
SensitiveData:S3Object/Credentials	Sensitive Data Identifications/Passwords/SensitiveData:S3Object-Credentials
SensitiveData:S3Object/CustomIdentifier	Sensitive Data Identifications/PII/SensitiveData:S3Object-CustomIdentifier
SensitiveData:S3Object/Financial	Sensitive Data Identifications/Financial/SensitiveData:S3Object-Financial
SensitiveData:S3Object/Multiple	Sensitive Data Identifications/PII/SensitiveData:S3Object-Multiple

Macie finding type	ASFF finding type
SensitiveData:S3Object/Personal	Sensitive Data Identifications/PII/SensitiveData:S3Object-Personal

Policy findings

If you configure Macie to publish policy findings to Security Hub, Macie automatically publishes each new policy finding that it creates and it does so immediately after it finishes processing the finding. If Macie detects a subsequent occurrence of an existing policy finding, it automatically publishes an update to the existing finding in Security Hub, using a publication frequency that you specify for your account. Macie performs these tasks for all policy findings that aren't archived automatically by a suppression rule.

If you're the Macie administrator for an organization, publication is limited to policy findings for S3 buckets that are owned directly by your account. Macie doesn't publish policy findings that it creates or updates for member accounts in your organization. This helps ensure that you don't have duplicate findings data in Security Hub.

As is the case for sensitive data findings, Macie uses the AWS Security Finding Format (ASFF) when it publishes new and updated policy findings to Security Hub. In the ASFF, the `Types` field uses a taxonomy that's slightly different from the finding type taxonomy in Macie.

The following table lists the ASFF finding type for each type of policy finding that Macie can create. If Macie created or updated a policy finding in Security Hub on or after January 28, 2021, the finding has one of the following values for the ASFF `Types` field in Security Hub.

Macie finding type	ASFF finding type
Policy:IAMUser/S3BlockPublicAccessDisabled	Software and Configuration Checks/AWS Security Best Practices/Policy:IAMUser-S3BlockPublicAccessDisabled
Policy:IAMUser/S3BucketEncryptionDisabled	Software and Configuration Checks/AWS Security Best Practices/Policy:IAMUser-S3BucketEncryptionDisabled

Macie finding type	ASFF finding type
Policy:IAMUser/S3BucketPublic	Effects/Data Exposure/Policy:IAMUser-S3BucketPublic
Policy:IAMUser/S3BucketReplicatedExternally	Software and Configuration Checks/AWS Security Best Practices/Policy:IAMUser-S3BucketReplicatedExternally
Policy:IAMUser/S3BucketSharedExternally	Software and Configuration Checks/AWS Security Best Practices/Policy:IAMUser-S3BucketSharedExternally
Policy:IAMUser/S3BucketSharedWithCloudFront	Software and Configuration Checks/AWS Security Best Practices/Policy:IAMUser-S3BucketSharedWithCloudFront

If Macie created or last updated a policy finding before January 28, 2021, the finding has one of the following values for the ASFF `Types` field in Security Hub:

- Policy:IAMUser/S3BlockPublicAccessDisabled
- Policy:IAMUser/S3BucketEncryptionDisabled
- Policy:IAMUser/S3BucketPublic
- Policy:IAMUser/S3BucketReplicatedExternally
- Policy:IAMUser/S3BucketSharedExternally

The values in the preceding list map directly to values for the **Finding type** (`type`) field in Macie.

Notes

As you review and process policy findings in Security Hub, note the following exceptions:

- In certain AWS Regions, Macie began using ASFF finding types for new and updated findings as early as January 25, 2021.

- If you acted upon a policy finding in Security Hub before Macie began using ASFF finding types in your AWS Region, the value for the ASFF `Types` field of the finding will be one of the Macie finding types in the preceding list. It will not be one of the ASFF finding types in the preceding table. This is true for policy findings that you acted upon using the AWS Security Hub console or the **BatchUpdateFindings** operation of the AWS Security Hub API.

Latency for publishing findings to Security Hub

When Amazon Macie creates a new policy or sensitive data finding, it publishes the finding to AWS Security Hub immediately after it finishes processing the finding.

If Macie detects a subsequent occurrence of an existing policy finding, it publishes an update to the existing Security Hub finding. The timing of the update depends on the publication frequency that you choose for your Macie account. By default, Macie publishes updates every 15 minutes. For more information, including how to change the setting for your account, see Configuring publication settings for findings.

Retrying publication when Security Hub isn't available

If AWS Security Hub isn't available, Amazon Macie creates a queue of findings that haven't been received by Security Hub. When the system is restored, Macie retries publication until the findings are received by Security Hub.

Updating existing findings in Security Hub

After Amazon Macie publishes a policy finding to AWS Security Hub, Macie updates the finding to reflect any additional occurrences of the finding or finding activity. Macie does this only for policy findings. Sensitive data findings, unlike policy findings, are all treated as new (unique).

When Macie publishes an update to a policy finding, Macie updates the value for the **Updated At** (`UpdatedAt`) field of the finding. You can use this value to determine when Macie most recently detected a subsequent occurrence of the potential policy violation or issue that produced the finding.

Macie might also update the value for the **Types** (`Types`) field of a finding if the existing value for the field isn't an ASFF finding type. This depends on whether you've acted upon the finding in Security Hub. If you haven't acted upon the finding, Macie changes the field's value to the appropriate ASFF finding type. If you've acted upon the finding, using either the AWS Security Hub

console or the **BatchUpdateFindings** operation of the AWS Security Hub API, Macie doesn't change the field's value.

Examples of Macie findings in AWS Security Hub

When Amazon Macie publishes findings to AWS Security Hub, it uses the AWS Security Finding Format (ASFF). This is the standard format for all findings in Security Hub. The following examples use sample data to demonstrate the structure and nature of the findings data that Macie publishes to Security Hub in this format:

- Example of a sensitive data finding
- Example of a policy finding

Example of a sensitive data finding in Security Hub

Here's an example of a sensitive data finding that Macie published to Security Hub using the ASFF.

```
{
    "SchemaVersion": "2018-10-08",
    "Id": "5be50fce24526e670df77bc00example",
    "ProductArn": "arn:aws:securityhub:us-east-1::product/aws/macie",
    "ProductName": "Macie",
    "CompanyName": "Amazon",
    "Region": "us-east-1",
    "GeneratorId": "aws/macie",
    "AwsAccountId": "111122223333",
    "Types":[
        "Sensitive Data Identifications/PII/SensitiveData:S3Object-Personal"
    ],
    "CreatedAt": "2022-05-11T10:23:49.667Z",
    "UpdatedAt": "2022-05-11T10:23:49.667Z",
    "Severity": {
        "Label": "HIGH",
        "Normalized": 70
    },
    "Title": "The S3 object contains personal information.",
    "Description": "The object contains personal information such as first or last
 names, addresses, or identification numbers.",
    "ProductFields": {
        "JobArn": "arn:aws:macie2:us-east-1:111122223333:classification-
job/698e99c283a255bb2c992feceexample",
```

```
        "S3Object.Path": "amzn-s3-demo-bucket/2022 Sourcing.tsv",
        "S3Object.Extension": "tsv",
        "S3Bucket.effectivePermission": "NOT_PUBLIC",
        "OriginType": "SENSITIVE_DATA_DISCOVERY_JOB",
        "S3Object.PublicAccess": "false",
        "S3Object.Size": "14",
        "S3Object.StorageClass": "STANDARD",
        "S3Bucket.allowsUnencryptedObjectUploads": "TRUE",
        "JobId": "698e99c283a255bb2c992feceexample",
        "aws/securityhub/FindingId": "arn:aws:securityhub:us-east-1::product/aws/
macie/5be50fce24526e670df77bc00example",
        "aws/securityhub/ProductName": "Macie",
        "aws/securityhub/CompanyName": "Amazon"
    },
    "Resources": [
        {
            "Type": "AwsS3Bucket",
            "Id": "arn:aws:s3:::amzn-s3-demo-bucket",
            "Partition": "aws",
            "Region": "us-east-1",
            "Details": {
                "AwsS3Bucket": {
                    "OwnerId":
 "7009a8971cd538e11f6b6606438875e7c86c5b672f46db45460ddcd08example",
                    "OwnerName": "johndoe",
                    "OwnerAccountId": "444455556666",
                    "CreatedAt": "2020-12-30T18:16:25.000Z",
                    "ServerSideEncryptionConfiguration": {
                        "Rules": [
                            {
                                "ApplyServerSideEncryptionByDefault": {
                                    "SSEAlgorithm": "aws:kms",
                                    "KMSMasterKeyID": "arn:aws:kms:us-
east-1:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab"
                                }
                            }
                        ]
                    },
                    "PublicAccessBlockConfiguration": {
                        "BlockPublicAcls": true,
                        "BlockPublicPolicy": true,
                        "IgnorePublicAcls": true,
                        "RestrictPublicBuckets": true
                    }
```

```
                }
            }
        },
        {
            "Type": "AwsS3Object",
            "Id": "arn:aws:s3:::amzn-s3-demo-bucket/2022 Sourcing.tsv",
            "Partition": "aws",
            "Region": "us-east-1",
            "DataClassification": {
                "DetailedResultsLocation": "s3://macie-data-discovery-results/
AWSLogs/111122223333/Macie/us-east-1/
                698e99c283a255bb2c992feceexample/111122223333/32b8485d-4f3a-3aa1-be33-
aa3f0example.jsonl.gz",
                "Result":{
                    "MimeType": "text/tsv",
                    "SizeClassified": 14,
                    "AdditionalOccurrences": false,
                    "Status": {
                        "Code": "COMPLETE"
                    },
                    "SensitiveData": [
                        {
                            "Category": "PERSONAL_INFORMATION",
                            "Detections": [
                                {
                                    "Count": 1,
                                    "Type": "USA_SOCIAL_SECURITY_NUMBER",
                                    "Occurrences": {
                                        "Cells": [
                                            {
                                                "Column": 10,
                                                "Row": 1,
                                                "ColumnName": "Other"
                                            }
                                        ]
                                    }
                                }
                            ],
                            "TotalCount": 1
                        }
                    ],
                    "CustomDataIdentifiers": {
                        "Detections": [
                        ],
```

```
                        "TotalCount": 0
                    }
                }
            },
            "Details": {
                "AwsS3Object": {
                    "LastModified": "2022-04-22T18:16:46.000Z",
                    "ETag": "ebe1ca03ee8d006d457444445example",
                    "VersionId": "SlBC72z5hArgexOJifxw_IN57example",
                    "ServerSideEncryption": "aws:kms",
                    "SSEKMSKeyId": "arn:aws:kms:us-
east-1:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab"
                }
            }
        }
    ],
    "WorkflowState": "NEW",
    "Workflow": {
        "Status": "NEW"
    },
    "RecordState": "ACTIVE",
    "FindingProviderFields": {
        "Severity": {
            "Label": "HIGH"
        },
        "Types": [
            "Sensitive Data Identifications/PII/SensitiveData:S3Object-Personal"
        ]
    },
    "Sample": false,
    "ProcessedAt": "2022-05-11T10:23:49.667Z"
}
```

Example of a policy finding in Security Hub

Here's an example of a new policy finding that Macie published to Security Hub in the ASFF.

```
{
    "SchemaVersion": "2018-10-08",
    "Id": "36ca8ba0-caf1-4fee-875c-37760example",
    "ProductArn": "arn:aws:securityhub:us-east-1::product/aws/macie",
    "ProductName": "Macie",
    "CompanyName": "Amazon",
```

```
    "Region": "us-east-1",
    "GeneratorId": "aws/macie",
    "AwsAccountId": "111122223333",
    "Types": [
        "Software and Configuration Checks/AWS Security Best Practices/Policy:IAMUser-
S3BlockPublicAccessDisabled"
    ],
    "CreatedAt": "2022-04-24T09:27:43.313Z",
    "UpdatedAt": "2022-04-24T09:27:43.313Z",
    "Severity": {
        "Label": "HIGH",
        "Normalized": 70
    },
    "Title": "Block Public Access settings are disabled for the S3 bucket",
    "Description": "All Amazon S3 block public access settings are disabled for the
 Amazon S3 bucket. Access to the bucket is
      controlled only by access control lists (ACLs) or bucket policies.",
    "ProductFields": {
        "S3Bucket.effectivePermission": "NOT_PUBLIC",
        "S3Bucket.allowsUnencryptedObjectUploads": "FALSE",
        "aws/securityhub/FindingId": "arn:aws:securityhub:us-east-1::product/aws/
macie/36ca8ba0-caf1-4fee-875c-37760example",
        "aws/securityhub/ProductName": "Macie",
        "aws/securityhub/CompanyName": "Amazon"
    },
    "Resources": [
        {
        "Type": "AwsS3Bucket",
        "Id": "arn:aws:s3:::amzn-s3-demo-bucket",
        "Partition": "aws",
        "Region": "us-east-1",
        "Tags": {
            "Team": "Recruiting",
            "Division": "HR"
        },
        "Details": {
            "AwsS3Bucket": {
              "OwnerId":
 "7009a8971cd538e11f6b6606438875e7c86c5b672f46db45460ddcd08example",
              "OwnerName": "johndoe",
              "OwnerAccountId": "444455556666",
              "CreatedAt": "2020-11-25T18:24:38.000Z",
              "ServerSideEncryptionConfiguration": {
                "Rules": [
```

```
                    {
                    "ApplyServerSideEncryptionByDefault": {
                        "SSEAlgorithm": "aws:kms",
                        "KMSMasterKeyID": "arn:aws:kms:us-
east-1:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab"
                    }
                  }
                ]
              },
              "PublicAccessBlockConfiguration": {
                "BlockPublicAcls": false,
                "BlockPublicPolicy": false,
                "IgnorePublicAcls": false,
                "RestrictPublicBuckets": false
               }
            }
         }
      }
    ],
    "WorkflowState": "NEW",
    "Workflow": {
        "Status": "NEW"
    },
    "RecordState": "ACTIVE",
    "FindingProviderFields": {
        "Severity": {
            "Label": "HIGH"
        },
        "Types": [
            "Software and Configuration Checks/AWS Security Best Practices/
Policy:IAMUser-S3BlockPublicAccessDisabled"
        ]
    },
    "Sample": false
}
```

Integrating Macie with AWS Security Hub

To integrate Amazon Macie with AWS Security Hub, enable Security Hub for your AWS account. To learn how, see Enabling Security Hub in the *AWS Security Hub User Guide*.

When you enable both Macie and Security Hub, the integration is enabled automatically. By default, Macie begins to publish new and updated policy findings to Security Hub automatically.

You don't need to take additional steps to configure the integration. If you have existing policy findings when the integration is enabled, Macie doesn't publish them to Security Hub. Instead, Macie publishes only those policy findings that it creates or updates after the integration is enabled.

You can optionally customize your configuration by choosing the frequency with which Macie publishes updates to policy findings in Security Hub. You can also choose to publish sensitive data findings to Security Hub. To learn how, see Configuring publication settings for findings.

Stopping publication of Macie findings to AWS Security Hub

To stop publishing Amazon Macie findings to AWS Security Hub, you can change the publication settings for your Macie account. To learn how, see Choosing publication destinations for findings. You can also do this by using Security Hub. To learn how, see Disabling the flow of findings from an integration in the *AWS Security Hub User Guide*.

Amazon EventBridge event schema for Macie findings

To support integration with other applications, services, and systems, such as monitoring or event management systems, Amazon Macie automatically publishes findings to Amazon EventBridge as events. EventBridge, formerly Amazon CloudWatch Events, is a serverless event bus service that delivers a stream of real-time data from applications and other AWS services to targets such as AWS Lambda functions, Amazon Simple Notification Service topics, and Amazon Kinesis streams. To learn more about EventBridge, see the Amazon EventBridge User Guide.

> **Note**
>
> If you currently use CloudWatch Events, note that EventBridge and CloudWatch Events are the same underlying service and API. However, EventBridge includes additional features that enable you to receive events from software as a service (SaaS) applications and your own applications. Because the underlying service and API are the same, the event schema for Macie findings is also the same.

Macie automatically publishes events for all new findings and subsequent occurrences of existing policy findings, except findings that are archived automatically by a suppression rule. The events are JSON objects that conform to the EventBridge schema for AWS events. Each event contains

a JSON representation of a particular finding. Because the data is structured as an EventBridge event, you can more easily monitor, process, and act upon a finding by using other applications, services, and tools. For details about how and when Macie publishes events for findings, see Configuring publication settings for findings.

Topics

Event schema for Macie findings

The following example shows the schema of an Amazon EventBridge event for an Amazon Macie finding. For detailed descriptions of the fields that can be included in a finding event, see Findings in the *Amazon Macie API Reference*. The structure and fields of a finding event map closely to the `Finding` object of the Amazon Macie API.

```
{
    "version": "0",
    "id": "event ID",
    "detail-type": "Macie Finding",
    "source": "aws.macie",
    "account": "AWS account ID (string)",
    "time": "event timestamp (string)",
    "region": "AWS Region (string)",
    "resources": [
        <-- ARNs of the resources involved in the event -->
    ],
    "detail": {
        <-- Details of a policy or sensitive data finding -->
    },
    "policyDetails": null, <-- Additional details of a policy finding or null for a
 sensitive data finding -->
    "sample": Boolean,
    "archived": Boolean
}
```

Example of an event for a policy finding

The following example uses sample data to demonstrate the structure and nature of objects and fields in an Amazon EventBridge event for a policy finding. In this example, the event reports a subsequent occurrence of an existing policy finding: Amazon Macie detected that block public access settings were disabled for an S3 bucket. The following fields and values can help you determine that this is the case:

- The `type` field is set to `Policy:IAMUser/S3BlockPublicAccessDisabled`.
- The `createdAt` and `updatedAt` fields have different values. This is one indicator that the event reports a subsequent occurrence of an existing policy finding. The values for these fields would be the same if the event reported a new finding.
- The `count` field is set to `2`, which indicates that this is the second occurrence of the finding.
- The `category` field is set to `POLICY`.
- The value for the `classificationDetails` field is `null`, which helps differentiate this event for a policy finding from an event for a sensitive data finding. For a sensitive data finding, this value would be a set of objects and fields that provide information about how and what sensitive data was found.

Also note that the value for the `sample` field is `true`. This value emphasizes that this is an example event for use in the documentation.

```
{
    "version": "0",
    "id": "0948ba87-d3b8-c6d4-f2da-732a1example",
    "detail-type": "Macie Finding",
    "source": "aws.macie",
    "account": "123456789012",
    "time": "2024-04-30T23:12:15Z",
    "region":"us-east-1",
    "resources": [],
    "detail": {
        "schemaVersion": "1.0",
        "id": "64b917aa-3843-014c-91d8-937ffexample",
        "accountId": "123456789012",
        "partition": "aws",
        "region": "us-east-1",
        "type": "Policy:IAMUser/S3BlockPublicAccessDisabled",
        "title": "Block public access settings are disabled for the S3 bucket",
```

```
        "description": "All bucket-level block public access settings were disabled for
 the S3 bucket. Access to the bucket is controlled by account-level block public access
 settings, access control lists (ACLs), and the bucket's bucket policy.",
        "severity": {
            "score": 3,
            "description": "High"
        },
        "createdAt": "2024-04-29T15:46:02Z",
        "updatedAt": "2024-04-30T23:12:15Z",
        "count": 2,
        "resourcesAffected": {
            "s3Bucket": {
                "arn": "arn:aws:s3:::amzn-s3-demo-bucket1",
                "name": "amzn-s3-demo-bucket1",
                "createdAt": "2020-04-03T20:46:56.000Z",
                "owner":{
                    "displayName": "johndoe",
                    "id":
 "7009a8971cd538e11f6b6606438875e7c86c5b672f46db45460ddcd08example"
                },
                "tags": [
                    {
                        "key": "Division",
                        "value": "HR"
                    },
                    {
                        "key": "Team",
                        "value": "Recruiting"
                    }
                ],
                "defaultServerSideEncryption": {
                    "encryptionType": "aws:kms",
                    "kmsMasterKeyId": "arn:aws:kms:us-
east-1:123456789012:key/1234abcd-12ab-34cd-56ef-1234567890ab"
                },
                "publicAccess": {
                    "permissionConfiguration": {
                        "bucketLevelPermissions": {
                            "accessControlList": {
                                "allowsPublicReadAccess": false,
                                "allowsPublicWriteAccess": false
                            },
                            "bucketPolicy": {
                                "allowsPublicReadAccess": false,
```

```
                                "allowsPublicWriteAccess": false
                            },
                            "blockPublicAccess": {
                                "ignorePublicAcls": false,
                                "restrictPublicBuckets": false,
                                "blockPublicAcls": false,
                                "blockPublicPolicy": false
                            }
                        },
                        "accountLevelPermissions": {
                            "blockPublicAccess": {
                                "ignorePublicAcls": true,
                                "restrictPublicBuckets": true,
                                "blockPublicAcls": true,
                                "blockPublicPolicy": true
                            }
                        }
                    },
                    "effectivePermission": "NOT_PUBLIC"
                },
                "allowsUnencryptedObjectUploads": "FALSE"
            },
            "s3Object": null
        },
        "category": "POLICY",
        "classificationDetails": null,
        "policyDetails": {
            "action": {
                "actionType": "AWS_API_CALL",
                "apiCallDetails": {
                    "api": "PutBucketPublicAccessBlock",
                    "apiServiceName": "s3.amazonaws.com",
                    "firstSeen": "2024-04-29T15:46:02.401Z",
                    "lastSeen": "2024-04-30T23:12:15.401Z"
                }
            },
            "actor": {
                "userIdentity": {
                    "type": "AssumedRole",
                    "assumedRole": {
                        "principalId": "AROA1234567890EXAMPLE:AssumedRoleSessionName",
                        "arn": "arn:aws:sts::123456789012:assumed-role/RoleToBeAssumed/
MySessionName",
                        "accountId": "111122223333",
```

```
                        "accessKeyId": "AKIAIOSFODNN7EXAMPLE",
                        "sessionContext": {
                            "attributes": {
                                "mfaAuthenticated": false,
                                "creationDate": "2024-04-29T10:25:43.511Z"
                            },
                            "sessionIssuer": {
                                "type": "Role",
                                "principalId": "AROA1234567890EXAMPLE",
                                "arn": "arn:aws:iam::123456789012:role/
RoleToBeAssumed",
                                "accountId": "123456789012",
                                "userName": "RoleToBeAssumed"
                            }
                        }
                    },
                    "root": null,
                    "iamUser": null,
                    "federatedUser": null,
                    "awsAccount": null,
                    "awsService": null
                },
                "ipAddressDetails":{
                    "ipAddressV4": "192.0.2.0",
                    "ipOwner": {
                        "asn": "-1",
                        "asnOrg": "ExampleFindingASNOrg",
                        "isp": "ExampleFindingISP",
                        "org": "ExampleFindingORG"
                    },
                    "ipCountry": {
                        "code": "US",
                        "name": "United States"
                    },
                    "ipCity": {
                        "name": "Ashburn"
                    },
                    "ipGeoLocation": {
                        "lat": 39.0481,
                        "lon": -77.4728
                    }
                },
                "domainDetails": null
            }
```

```
        },
        "sample": true,
        "archived": false
    }
}
```

Example of an event for a sensitive data finding

The following example uses sample data to demonstrate the structure and nature of objects and fields in an Amazon EventBridge event for a sensitive data finding. In this example, the event reports a new sensitive data finding: Amazon Macie found multiple categories and types of sensitive data in an S3 object. The following fields and values can you help you determine that this is the case:

- The `type` field is set to `SensitiveData:S3Object/Multiple`.
- The `createdAt` and `updatedAt` fields have the same values. Unlike policy findings, this is always the case for sensitive data findings. All sensitive data findings are considered new.
- The `count` field is set to 1, which indicates that this is a new finding. Unlike policy findings, this is always the case for sensitive data findings. All sensitive data findings are considered unique (new).
- The `category` field is set to `CLASSIFICATION`.
- The value for the `policyDetails` field is `null`, which helps differentiate this event for a sensitive data finding from an event for a policy finding. For a policy finding, this value would be a set of objects and fields that provide information about a potential policy violation or issue with the security or privacy of an S3 bucket.

Also note that the value for the `sample` field is `true`. This value emphasizes that this is an example event for use in the documentation.

```
{
    "version": "0",
    "id": "14ddd0b1-7c90-b9e3-8a68-6a408example",
    "detail-type": "Macie Finding",
    "source": "aws.macie",
    "account": "123456789012",
    "time": "2024-04-20T08:19:10Z",
    "region": "us-east-1",
    "resources": [],
```

```
    "detail": {
        "schemaVersion": "1.0",
        "id": "4ed45d06-c9b9-4506-ab7f-18a57example",
        "accountId": "123456789012",
        "partition": "aws",
        "region": "us-east-1",
        "type": "SensitiveData:S3Object/Multiple",
        "title": "The S3 object contains multiple categories of sensitive data",
        "description": "The S3 object contains more than one category of sensitive
 data.",
        "severity": {
            "score": 3,
            "description": "High"
        },
        "createdAt": "2024-04-20T18:19:10Z",
        "updatedAt": "2024-04-20T18:19:10Z",
        "count": 1,
        "resourcesAffected": {
            "s3Bucket": {
                "arn": "arn:aws:s3:::amzn-s3-demo-bucket2",
                "name": "amzn-s3-demo-bucket2",
                "createdAt": "2020-05-15T20:46:56.000Z",
                "owner": {
                    "displayName": "johndoe",
                    "id":
 "7009a8971cd538e11f6b6606438875e7c86c5b672f46db45460ddcd08example"
                },
                "tags":[
                    {
                        "key":"Division",
                        "value":"HR"
                    },
                    {
                        "key":"Team",
                        "value":"Recruiting"
                    }
                ],
                "defaultServerSideEncryption": {
                    "encryptionType": "aws:kms",
                    "kmsMasterKeyId": "arn:aws:kms:us-
east-1:123456789012:key/1234abcd-12ab-34cd-56ef-1234567890ab"
                },
                "publicAccess": {
                    "permissionConfiguration": {
```

```
                        "bucketLevelPermissions": {
                            "accessControlList": {
                                "allowsPublicReadAccess": false,
                                "allowsPublicWriteAccess": false
                            },
                            "bucketPolicy":{
                                "allowsPublicReadAccess": false,
                                "allowsPublicWriteAccess": false
                            },
                            "blockPublicAccess": {
                                "ignorePublicAcls": true,
                                "restrictPublicBuckets": true,
                                "blockPublicAcls": true,
                                "blockPublicPolicy": true
                            }
                        },
                        "accountLevelPermissions": {
                            "blockPublicAccess": {
                                "ignorePublicAcls": false,
                                "restrictPublicBuckets": false,
                                "blockPublicAcls": false,
                                "blockPublicPolicy": false
                            }
                        }
                    },
                    "effectivePermission": "NOT_PUBLIC"
                },
                "allowsUnencryptedObjectUploads": "TRUE"
            },
            "s3Object":{
                "bucketArn": "arn:aws:s3:::amzn-s3-demo-bucket2",
                "key": "2024 Sourcing.csv",
                "path": "amzn-s3-demo-bucket2/2024 Sourcing.csv",
                "extension": "csv",
                "lastModified": "2024-04-19T22:08:25.000Z",
                "versionId": "",
                "serverSideEncryption": {
                    "encryptionType": "aws:kms",
                    "kmsMasterKeyId": "arn:aws:kms:us-
east-1:123456789012:key/1234abcd-12ab-34cd-56ef-1234567890ab"
                },
                "size": 4750,
                "storageClass": "STANDARD",
                "tags":[
```

```
                    {
                        "key":"Division",
                        "value":"HR"
                    },
                    {
                        "key":"Team",
                        "value":"Recruiting"
                    }
                ],
                "publicAccess": false,
                "etag": "6bb7fd4fa9d36d6b8fb8882caexample"
            }
        },
        "category": "CLASSIFICATION",
        "classificationDetails": {
            "jobArn": "arn:aws:macie2:us-east-1:123456789012:classification-
job/3ce05dbb7ec5505def334104bexample",
            "jobId": "3ce05dbb7ec5505def334104bexample",
            "result": {
                "status": {
                    "code": "COMPLETE",
                    "reason": null
                },
                "sizeClassified": 4750,
                "mimeType": "text/csv",
                "additionalOccurrences": true,
                "sensitiveData": [
                    {
                        "category": "PERSONAL_INFORMATION",
                        "totalCount": 65,
                        "detections": [
                            {
                                "type": "USA_SOCIAL_SECURITY_NUMBER",
                                "count": 30,
                                "occurrences": {
                                    "lineRanges": null,
                                    "offsetRanges": null,
                                    "pages": null,
                                    "records": null,
                                    "cells": [
                                        {
                                            "row": 2,
                                            "column": 1,
                                            "columnName": "SSN",
```

```
                                    "cellReference": null
                                },
                                {
                                    "row": 3,
                                    "column": 1,
                                    "columnName": "SSN",
                                    "cellReference": null
                                },
                                {
                                    "row": 4,
                                    "column": 1,
                                    "columnName": "SSN",
                                    "cellReference": null
                                }
                            ]
                        }
                    },
                    {
                        "type": "NAME",
                        "count": 35,
                        "occurrences": {
                            "lineRanges": null,
                            "offsetRanges": null,
                            "pages": null,
                            "records": null,
                            "cells": [
                                {
                                    "row": 2,
                                    "column": 3,
                                    "columnName": "Name",
                                    "cellReference": null
                                },
                                {
                                    "row": 3,
                                    "column": 3,
                                    "columnName": "Name",
                                    "cellReference": null
                                }
                            ]
                        }
                    }
                ]
            },
            {
```

```
                        "category": "FINANCIAL_INFORMATION",
                        "totalCount": 30,
                        "detections": [
                            {
                                "type": "CREDIT_CARD_NUMBER",
                                "count": 30,
                                "occurrences": {
                                    "lineRanges": null,
                                    "offsetRanges": null,
                                    "pages": null,
                                    "records": null,
                                    "cells": [
                                        {
                                            "row": 2,
                                            "column": 14,
                                            "columnName": "CCN",
                                            "cellReference": null
                                        },
                                        {
                                            "row": 3,
                                            "column": 14,
                                            "columnName": "CCN",
                                            "cellReference": null
                                        }
                                    ]
                                }
                            }
                        ]
                    }
                ],
                "customDataIdentifiers": {
                    "totalCount": 0,
                    "detections": []
                }
            },
            "detailedResultsLocation": "s3://macie-data-discovery-results/
AWSLogs/123456789012/Macie/us-east-1/3ce05dbb7ec5505def334104bexample/
d48bf16d-0deb-3e49-9d8c-d407cexample.jsonl.gz",
            "originType": "SENSITIVE_DATA_DISCOVERY_JOB"
        },
        "policyDetails": null,
        "sample": true,
        "archived": false
    }
```

```
}
```

Forecasting and monitoring Macie costs

To help you forecast and monitor your costs for using Amazon Macie, Macie calculates and provides estimated usage costs for your account. With this data, you can determine whether to adjust your use of the service or your account quotas. If you're currently participating in a 30-day free trial of Macie, you can use this data to estimate your costs for using Macie after the free trial ends. You can also check the status of your trial.

You can review your estimated usage costs on the Amazon Macie console and access them programmatically with the Amazon Macie API. If you're the Macie administrator for an organization, you can review and access both aggregated data for your organization and breakdowns of the data for accounts in your organization.

In addition to the estimated usage costs that Macie provides, you can review and monitor your actual costs by using AWS Billing and Cost Management. AWS Billing and Cost Management provides features that are designed to help you track and analyze your costs for AWS services, and manage budgets for your account or organization. It also provides features that can help you forecast usage costs based on historical data. To learn more, see the AWS Billing User Guide.

Topics

Understanding estimated usage costs for Macie

Amazon Macie pricing is based on the following dimensions.

Preventative control monitoring

These costs derive from maintaining an inventory of your Amazon Simple Storage Service (Amazon S3) general purpose buckets, and evaluating and monitoring the buckets for security and access control. For more information, see How Macie monitors Amazon S3 data security.

You're charged based on the total number of S3 general purpose buckets that Macie evaluates and monitors for your account, for up to 10,000 buckets. The charges are prorated per day.

Object monitoring for automated sensitive data discovery

These costs derive from monitoring and evaluating your S3 bucket inventory to identify S3 objects that are eligible for analysis by automated sensitive data discovery. For more information, see How automated sensitive data discovery works.

You're charged based on the total number of S3 objects that are stored in general purpose buckets for your account. The charges are prorated per day.

Object analysis by sensitive data discovery jobs and automated sensitive data discovery

These costs derive from analyzing S3 objects and reporting sensitive data that Macie finds in the objects. This includes analyses and reporting by sensitive data discovery jobs and by automated sensitive data discovery. For more information, see Discovering sensitive data.

You're charged based on the amount of uncompressed data that Macie analyzes in S3 objects. There's no charge for objects that Macie can't analyze for reasons such as use of an unsupported Amazon S3 storage class, use of an unsupported file or storage format, or permissions settings. In addition, these costs don't vary based on the number of sensitive data findings produced by your jobs or by automated sensitive data discovery.

To manage costs for automated sensitive data discovery, you can exclude individual S3 buckets from the analyses. For example, you might exclude buckets that are known to meet your organization's security and compliance requirements. If your account is part of an organization that centrally manages multiple Macie accounts, an additional option is to selectively enable or disable automated sensitive data discovery for individual accounts in your organization. For more information, see Configuring settings for automated sensitive data discovery.

Costs for sensitive data discovery jobs are restricted by the monthly sensitive data discovery quota for your account. (The default quota is 5 TB of data.) If a job is running and the analysis of eligible objects reaches this quota, Macie automatically pauses the job until the next calendar month starts and the monthly quota is reset for your account, or you increase the quota for your account.

If you're the Macie administrator for an organization, costs for sensitive data discovery jobs are restricted by the monthly sensitive data discovery quota for each account that you analyze data for. The quota for a member account defines the maximum amount of data that your jobs and the member account's jobs can analyze for the account during a calendar month. If a job is running and the analysis of eligible objects reaches this quota for a member account, Macie stops analyzing objects in buckets that the account owns. When Macie finishes analyzing

objects for all other accounts that haven't met the quota, Macie automatically pauses the job. If it's a one-time job, Macie automatically resumes the job when the next calendar month starts or the quota is increased for all the affected accounts, whichever occurs first. If it's a periodic job, Macie automatically resumes the job when the next run is scheduled to start or the next calendar month starts, whichever occurs first. If a scheduled run starts before the next calendar month starts or the quota is increased for an affected account, Macie doesn't analyze objects in buckets that the account owns.

> **Tip**
>
> For helpful tips about managing or reducing sensitive data discovery costs, see the following blog post on the *AWS Security Blog*: How to use Amazon Macie to reduce the cost of discovering sensitive data.

For detailed information and examples of usage costs, see Amazon Macie pricing.

When you use Macie to review your estimated usage costs, it's important to understand how the cost estimates are calculated. Consider the following:

- The estimates are reported in US dollars (USD) and are for the current AWS Region only. If you use Macie in multiple Regions, the data isn't aggregated for all the Regions in which you use Macie.
- On the console, the estimates are inclusive for the current calendar month to date. If you query the data programmatically with the Amazon Macie API, you can choose an inclusive time range for the estimates. This can be a rolling time range of the preceding 30 days or the current calendar month to date.
- The estimates don't reflect all the discounts that might apply to your account. The exception is discounts that derive from Regional volume pricing tiers, as described in Amazon Macie pricing. If your account qualifies for this type of discount, the estimates reflect that discount.
- If you're the Macie administrator for an organization, the estimates don't reflect combined usage volume discounts for your organization. For information about these discounts, see Volume discounts in the *AWS Billing User Guide*.
- For preventative control monitoring, the estimate is based on the average daily cost for the applicable time range. The cost is prorated per day.
- For automated sensitive data discovery, the overall estimate is based on the average daily cost for object monitoring (prorated per day) and the amount of uncompressed data that Macie

has analyzed thus far during the applicable time range. If you're the Macie administrator for an organization and you enable automated sensitive data discovery for member accounts, the estimated costs of those activities are included in the estimates for each applicable member account.

- For sensitive data discovery jobs, the estimate is based on the amount of uncompressed data that your jobs have analyzed thus far during the applicable time range. If you're the Macie administrator for an organization and you run jobs that analyze data for member accounts, the estimated costs of those jobs are included in the estimate for each applicable member account.
- If your account is a member account in an organization and your Macie administrator enables automated sensitive data discovery or runs sensitive data discovery jobs that analyze your data, the estimated costs of those activities are included in the estimates for your account.
- The estimates don't include costs that you incur for using other AWS services with certain Macie features. For example, using customer managed AWS KMS keys to decrypt S3 objects that you want to inspect for sensitive data.

Also note that Macie provides a monthly free tier for analysis of S3 objects by sensitive data discovery jobs and automated sensitive data discovery. Each month, there's no charge to analyze up to 1 GB of data to discover and report sensitive data in S3 objects. If more than 1 GB of data is analyzed during a given month, sensitive data discovery charges begin to accrue for your account after the first 1 GB of data. If less than 1 GB of data is analyzed during a given month, the remaining allocation doesn't roll over to the next month. If your account is part of an organization with consolidated billing, the free tier applies to the combined amount of data analyzed for your organization. In other words, there's no charge to analyze up to 1 GB of data each month for all the accounts in your organization.

Reviewing estimated usage costs for Macie

To review your current estimated usage costs for Amazon Macie, you can use the Amazon Macie console or the Amazon Macie API. Both the console and the API provide estimated costs for Macie pricing dimensions. If you're currently participating in a 30-day free trial, you can use this data to estimate your costs for using Macie after your free trial ends. For information about Macie pricing dimensions and considerations, see Understanding estimated usage costs. For detailed information and examples of usage costs, see Amazon Macie pricing.

In Macie, estimated usage costs are reported in US dollars (USD) and apply only to the current AWS Region. If you use the console to review the data, the cost estimates are for the current

calendar month to date (inclusively). If you query the data programmatically with the Amazon Macie API, you can specify an inclusive time range for the estimates, either a rolling time range of the preceding 30 days or the current calendar month to date.

Topics

Reviewing estimated usage costs on the Amazon Macie console

On the Amazon Macie console, cost estimates are organized as follows:

- **Preventative control monitoring** – This is the estimated cost of maintaining an inventory of your Amazon Simple Storage Service (Amazon S3) general purpose buckets, and evaluating and monitoring the buckets for security and access control.
- **Sensitive data discovery jobs** – This is the estimated cost of the sensitive data discovery jobs that you ran.
- **Automated sensitive data discovery** – These are the estimated costs of performing automated sensitive data discovery. This includes monitoring and evaluating your S3 bucket inventory to identify S3 objects that are eligible for analysis. It also includes analyzing eligible objects and reporting sensitive data statistics, findings, and other types of results.

To review estimates for automated sensitive data discovery by using the console, you must be the Macie administrator for an organization or have a standalone Macie account.

To review your estimated usage costs on the console

Follow these steps to review your estimated usage costs by using the Amazon Macie console.

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to review your estimated costs.
3. In the navigation pane, choose **Usage**.

If you have a standalone Macie account or a member account in an organization, the **Usage** page displays a breakdown of the estimated usage costs for your account.

If you're the Macie administrator for an organization, the **Usage** page lists accounts in your organization. In the table:

- **Service quota – Jobs** – This is the current monthly quota for running sensitive data discovery jobs to analyze S3 objects in buckets that an account owns.
- **Free trial** – These fields indicate whether an account is currently participating in the free trial for preventative control monitoring or automated sensitive data discovery. A **Free trial** field is empty if the applicable free trial has ended for an account.
- **Total** – This is the total estimated cost for an account.

The **Estimated costs** section shows the total estimated cost for your organization and a breakdown of those costs. To review the breakdown of estimated costs for a specific account in your organization, choose the account in the table. The **Estimated costs** section then shows this breakdown. To show this data for another account, choose the account in the table. To clear your account selection, choose **X** next to the account ID.

Querying estimated usage costs with the Amazon Macie API

To query your estimated usage costs programmatically, you can use the following operations of the Amazon Macie API:

- **GetUsageTotals** – This operation returns total estimated usage costs for your account, grouped by usage metric. If you're the Macie administrator for an organization, this operation returns aggregated cost estimates for all the accounts in your organization. To learn more about this operation, see Usage Totals in the *Amazon Macie API Reference*.
- **GetUsageStatistics** – This operation returns usage statistics and related data for your account, grouped by account and then by usage metric. The data includes total estimated usage costs and current account quotas. As applicable, it also indicates when your 30-day free trial started for Macie and for automated sensitive data discovery. If you're the Macie administrator for an organization, this operation returns a breakdown of the data for all the accounts in your organization. You can customize your query by sorting and filtering the query results. To learn more about this operation, see Usage Statistics in the *Amazon Macie API Reference*.

When you use either operation, you can optionally specify an inclusive time range for the data. This time range can be a rolling time range of the preceding 30 days (`PAST_30_DAYS`) or the current

calendar month to date (MONTH_TO_DATE). If you don't specify a time range, Macie returns the data for the preceding 30 days.

The following examples show how to query estimated usage costs and statistics by using the AWS Command Line Interface (AWS CLI). You can also query the data by using a current version of another AWS command line tool or an AWS SDK, or by sending HTTPS requests directly to Macie. For information about AWS tools and SDKs, see Tools to Build on AWS.

Examples

- Example 1: Querying total estimated usage costs
- Example 2: Querying usage statistics

Example 1: Querying total estimated usage costs

To query total estimated usage costs by using the AWS CLI, run the get-usage-totals command and optionally specify a time range for the data. For example:

```
C:\> aws macie2 get-usage-totals --time-range MONTH_TO_DATE
```

Where *MONTH_TO_DATE* specifies the current calendar month to date as the time range for the data.

If the command runs successfully, you receive output similar to the following.

```
{
    "timeRange": "MONTH_TO_DATE",
    "usageTotals": [
        {
            "currency": "USD",
            "estimatedCost": "153.45",
            "type": "SENSITIVE_DATA_DISCOVERY"
        },
        {
            "currency": "USD",
            "estimatedCost": "65.18",
            "type": "AUTOMATED_SENSITIVE_DATA_DISCOVERY"
        },
        {
            "currency": "USD",
            "estimatedCost": "1.51",
```

```
            "type": "DATA_INVENTORY_EVALUATION"
        },
        {
            "currency": "USD",
            "estimatedCost": "0.98",
            "type": "AUTOMATED_OBJECT_MONITORING"
        }
    ]
}
```

Where `estimatedCost` is the total estimated usage cost for the associated usage metric (`type`):

- `SENSITIVE_DATA_DISCOVERY`, for analyzing S3 objects with sensitive data discovery jobs.
- `AUTOMATED_SENSITIVE_DATA_DISCOVERY`, for analyzing S3 objects with automated sensitive data discovery.
- `DATA_INVENTORY_EVALUATION`, for monitoring and evaluating S3 general purpose buckets for security and access control.
- `AUTOMATED_OBJECT_MONITORING`, for evaluating and monitoring your S3 bucket inventory to identify S3 objects that are eligible for analysis by automated sensitive data discovery.

Example 2: Querying usage statistics

To query usage statistics by using the AWS CLI, run the get-usage-statistics command. You can optionally sort, filter, and specify a time range for the query results. The following example retrieves usage statistics for a Macie administrator account for the preceding 30 days. The results are sorted in ascending order by AWS account ID.

For Linux, macOS, or Unix, using the backslash (\) line-continuation character to improve readability:

```
$ aws macie2 get-usage-statistics \
--sort-by '{"key":"accountId","orderBy":"ASC"}' \
--time-range PAST_30_DAYS
```

For Microsoft Windows, using the caret (^) line-continuation character to improve readability:

```
C:\> aws macie2 get-usage-statistics ^
--sort-by={\"key\":\"accountId\",\"orderBy\":\"ASC\"} ^
```

```
--time-range PAST_30_DAYS
```

Where:

- *accountId* specifies the field to use to sort the results.
- *ASC* is the sort order to apply to the results, based on the value for the specified field (*accountId*).
- *PAST_30_DAYS* specifies the preceding 30 days as the time range for the data.

If the command runs successfully, Macie returns a `records` array. The array contains an object for each account that's included in the query results. For example:


```
{
    "records": [
        {
            "accountId": "111122223333",
            "automatedDiscoveryFreeTrialStartDate": "2024-01-28T16:00:00+00:00",
            "freeTrialStartDate": "2020-05-20T12:26:36.917000+00:00",
            "usage": [
                {
                    "currency": "USD",
                    "estimatedCost": "1.51",
                    "type": "DATA_INVENTORY_EVALUATION"
                },
                {
                    "currency": "USD",
                    "estimatedCost": "65.18",
                    "type": "AUTOMATED_SENSITIVE_DATA_DISCOVERY"
                },
                {
                    "currency": "USD",
                    "estimatedCost": "153.45",
                    "serviceLimit": {
                        "isServiceLimited": false,
                        "unit": "TERABYTES",
                        "value": 50
                    },
                    "type": "SENSITIVE_DATA_DISCOVERY"
                },
                {
                    "currency": "USD",
```

```
                    "estimatedCost": "0.98",
                    "type": "AUTOMATED_OBJECT_MONITORING"
                }
            ]
        },
        {
            "accountId": "444455556666",
            "automatedDiscoveryFreeTrialStartDate": "2024-01-28T16:00:00+00:00",
            "freeTrialStartDate": "2020-05-18T16:26:36.917000+00:00",
            "usage": [
                {
                    "currency": "USD",
                    "estimatedCost": "1.58",
                    "type": "DATA_INVENTORY_EVALUATION"
                },
                {
                    "currency": "USD",
                    "estimatedCost": "63.13",
                    "type": "AUTOMATED_SENSITIVE_DATA_DISCOVERY"
                },
                {
                    "currency": "USD",
                    "estimatedCost": "145.12",
                    "serviceLimit": {
                        "isServiceLimited": false,
                        "unit": "TERABYTES",
                        "value": 50
                    },
                    "type": "SENSITIVE_DATA_DISCOVERY"
                },
                {
                    "currency": "USD",
                    "estimatedCost": "1.02",
                    "type": "AUTOMATED_OBJECT_MONITORING"
                }
            ]
        }
    ],
    "timeRange": "PAST_30_DAYS"
}
```


Where `estimatedCost` is the total estimated usage cost for the associated usage metric (`type`) for an account:

- `DATA_INVENTORY_EVALUATION`, for monitoring and evaluating S3 general purpose buckets for security and access control.
- `AUTOMATED_SENSITIVE_DATA_DISCOVERY`, for analyzing S3 objects with automated sensitive data discovery.
- `SENSITIVE_DATA_DISCOVERY`, for analyzing S3 objects with sensitive data discovery jobs.
- `AUTOMATED_OBJECT_MONITORING`, for evaluating and monitoring the account's S3 bucket inventory to identify S3 objects that are eligible for analysis by automated sensitive data discovery.

Participating in the free trial of Macie

When you enable Amazon Macie for the first time, your AWS account is automatically enrolled in the 30-day free trial of Macie. This includes individual member accounts in an AWS Organizations organization.

During the free trial, there's no charge for using Macie in a specific AWS Region to:

- **Perform preventative control monitoring** – This includes generating and maintaining an inventory of your Amazon Simple Storage Service (Amazon S3) general purpose buckets in the Region. It also includes evaluating and monitoring the buckets for security and access control.

  For more information, see How Macie monitors Amazon S3 data security.
- **Perform automated sensitive data discovery** – This includes monitoring and evaluating your S3 bucket inventory in the Region to identify S3 objects that are eligible for analysis. It also includes analyzing eligible objects and reporting sensitive data statistics, findings, and other types of results. To configure and manage this feature, you must be the Macie administrator for an organization or have a standalone Macie account. If you're a Macie administrator, you can use this feature to analyze objects in S3 buckets that your member accounts own.

  For more information, see How automated sensitive data discovery works.

For a list of Regions where Macie is currently available, see Amazon Macie endpoints and quotas in the *AWS General Reference*.

The free trial runs for 30 consecutive days. You can't pause it after it starts. After the free trial ends, charges begin to accrue for performing preventative control monitoring. Charges also begin to accrue for performing automated sensitive data discovery. If you're the Macie administrator for an

organization, charges accrue as applicable for each account in your organization. You can use Macie to review breakdowns of estimated usage costs for individual accounts in your organization.

Notes

During the free trial, you might incur charges for other AWS services that you use with certain Macie features—for example, using customer managed AWS KMS keys to decrypt S3 objects that you want to inspect for sensitive data.
The free trial doesn't include analysis of S3 objects by sensitive data discovery jobs. You'll incur charges if you create and run sensitive data discovery jobs that analyze more than 1 GB of uncompressed data during the free trial. (Macie provides a monthly free tier for sensitive data discovery. Each month, there's no charge to analyze up to 1 GB of uncompressed data in S3 objects. After the first 1 GB of data, costs accrue.)

During the free trial, you can check the status of your trial and review estimated usage costs for your account. The cost estimates are based on your use of Macie thus far during the free trial. They can help you understand what some of your usage costs might be after the trial ends. For details about how Macie calculates these values, see Understanding estimated usage costs.

To check your status and estimated costs during the free trial

Follow these steps to check the status of your trial and review your estimated usage costs by using the Amazon Macie console. To access this data programmatically, you can use the GetUsageStatistics operation of the Amazon Macie API.

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to check the status of your free trial and your estimated usage costs.
3. In the navigation pane, choose **Usage**.

The **Usage** page indicates the number of remaining days in your free trial. It also shows a breakdown of your estimated usage costs in US dollars (USD):

- **Preventative control monitoring** – This is the total projected cost of maintaining an inventory of your S3 general purpose buckets, and evaluating and monitoring the buckets for security and access control after the free trial ends.

- **Sensitive data discovery jobs** – This is the total estimated cost of any sensitive data discovery jobs that you ran. Sensitive data discovery jobs aren't included in the free trial.
- **Automated sensitive data discovery** – These are the total projected costs of performing automated sensitive data discovery after the free trial ends, broken down by pricing dimension—object monitoring and object analysis. To review these estimates on the console, you must be the Macie administrator for an organization or have a standalone Macie account.

If you're the Macie administrator for an organization, the **Usage** page provides details about the accounts in your organization. In the table:

- **Service quota – Jobs** – This is the current monthly quota for running sensitive data discovery jobs to analyze S3 objects in buckets that an account owns.
- **Free trial** – These fields indicate whether an account is currently participating in the free trial for preventative control monitoring or automated sensitive data discovery. A **Free trial** field is empty if the applicable free trial has ended for an account.
- **Total** – This is the total estimated cost for an account.

The **Estimated costs** section shows estimated costs for your organization overall. To review the breakdown of estimated costs for a specific account in your organization, choose the account in the table. The **Estimated costs** section then shows this breakdown. To show this data for another account, choose the account in the table. To clear your account selection, choose **X** next to the account ID.

Notes

If an account stores more than 150 TB of data in Amazon S3, the account's estimated and actual costs for automated sensitive data discovery might be higher than the cost projections that Macie provides during the 30-day free trial. This is because object analysis by automated sensitive data discovery is paused when 150 GB of uncompressed data has been analyzed for an account that's enrolled in the free trial. Object analysis resumes for the account after the free trial ends. For assistance forecasting costs for an account that stores more than 150 TB of data in Amazon S3, contact AWS Support.

To manage costs for automated sensitive data discovery after the free trial ends, you can exclude individual S3 buckets from subsequent analyses. If you're the Macie administrator for an organization, an additional option is to selectively enable or disable automated

sensitive data discovery for individual accounts in your organization. For information about these options, see Configuring settings for automated sensitive data discovery.

Managing multiple Macie accounts as an organization

If your AWS environment has multiple accounts, you can associate the Amazon Macie accounts in your environment and centrally manage them as an organization in Macie. With this configuration, a designated Macie administrator can assess and monitor the overall security posture of your organization's Amazon Simple Storage Service (Amazon S3) data estate, and discover sensitive data in your organization's S3 buckets. The administrator can also perform various account management and administration tasks at scale, such as monitoring estimated usage costs and assessing account quotas.

In Macie, an organization consists of a designated Macie administrator account and one or more associated member accounts. You can associate the accounts in two ways, by integrating Macie with AWS Organizations or by sending and accepting membership invitations in Macie. We recommend that you integrate Macie with AWS Organizations.

AWS Organizations is a global account management service that enables AWS administrators to consolidate and centrally manage multiple AWS accounts. It provides account management and consolidated billing features that are designed to support budgetary, security, and compliance needs. It's offered at no additional charge and it integrates with multiple AWS services, including Macie, AWS Security Hub, and Amazon GuardDuty. To learn more, see the AWS Organizations User Guide.

If you prefer to centrally manage multiple Macie accounts without using AWS Organizations, you can use membership invitations instead. If you send an invitation and it's accepted by another account, your account becomes the Macie administrator account for the other account. If you receive and accept an invitation, your account becomes a Macie member account and the Macie administrator account can access and manage certain settings, data, and resources for your Macie account.

Topics

Macie administrator and member account relationships

If you centrally manage multiple Amazon Macie accounts as an organization, the Macie administrator has access to Amazon Simple Storage Service (Amazon S3) inventory data, policy findings, and certain Macie settings and resources for associated member accounts. The administrator can also enable automated sensitive data discovery and run sensitive data discovery jobs to detect sensitive data in S3 buckets that member accounts own. Support for specific tasks varies based on whether a Macie administrator account is associated with a member account through AWS Organizations or by invitation.

The following table provides details about the relationship between Macie administrator and member accounts. It indicates the default permissions for each type of account. To further restrict access to Macie features and operations, you can use custom AWS Identity and Access Management (IAM) policies.

In the table:

- **Self** indicates that the account can't perform the task for any associated accounts.
- **Any** indicates that the account can perform the task for an individual associated account.
- **All** indicates that the account can perform the task and the task applies to all associated accounts.

A dash (–) indicates that the account can’t perform the task.

Task	Through AWS Organizations		By invitation	
	Administrator	**Member**	**Administrator**	**Member**
Enable Macie	Any	–	Self	Self
Review the organization's account inventory [1]	All	–	All	–
Add a member account	Any	–	Any	–

Review statistics and metadata for S3 buckets	All	Self	All	Self
Review policy findings	All	Self	All	Self
Suppress (archive) policy findings [2]	All	–	All	–
Publish policy findings [3]	Self	Self	Self	Self
Configure a repository for sensitive data discovery results [4]	Self	Self	Self	Self
Create and use allow lists	Self	Self	Self	Self
Create and use custom data identifiers	Self	Self	Self	Self
Configure automated sensitive data discovery settings	All	–	All	–
Enable or disable automated sensitive data discovery	Any	–	Any	–

Review automated sensitive data discovery statistics, data, and results [5]	All	Self	All	Self
Create and run sensitive data discovery jobs [6]	Any	Self	Any	Self
Review the details of sensitive data discovery jobs [7]	Self	Self	Self	Self
Review sensitive data findings [8]	Self	Self	Self	Self
Suppress (archive) sensitive data findings [8]	Self	Self	Self	Self
Publish sensitive data findings [8]	Self	Self	Self	Self
Configure Macie to retrieve sensitive data samples for findings	Self	Self	Self	Self
Retrieve sensitive data samples for findings [9]	Self	Self	Self	Self

Configure publication destinations for findings	Self	Self	Self	Self
Set the publication frequency for findings	All	Self	All	Self
Create sample findings	Self	Self	Self	Self
Review account quotas and estimated usage costs	All	Self	All	Self
Suspend Macie [10]	Any	–	Any	Self
Disable Macie [11]	Self	Self	Self	Self
Remove (disassociate) a member account	Any	–	Any	–
Disassociate from an administrator account	–	–	–	Self
Delete an association with another account [12]	Any	–	Any	Self

1.

The administrator for an organization in AWS Organizations can review all accounts in the organization, including accounts that haven't enabled Macie. The administrator for an invitation-based organization can review only those accounts that they add to their inventory.

2. Only an administrator can suppress policy findings. If an administrator creates a suppression rule, Macie applies the rule to policy findings for all accounts in the organization unless the rule is configured to exclude specific accounts. If a member creates a suppression rule, Macie doesn't apply the rule to policy findings for the member's account.

3. Only the account that owns an affected resource can publish policy findings for the resource to AWS Security Hub. Both administrator and member accounts automatically publish policy findings for an affected resource to Amazon EventBridge.

4. If an administrator enables automated sensitive data discovery or configures a job to analyze objects in S3 buckets that a member account owns, Macie stores the sensitive data discovery results in the repository for the administrator account.

5. Only an administrator can access sensitive data findings that automated sensitive data discovery produces. Both an administrator and a member can review other types of data that automated sensitive data discovery produces for the member's account.

6. A member can configure a job to analyze objects only in S3 buckets that their account owns. An administrator can configure a job to analyze objects in buckets that their account owns or a member account owns. For information about how quotas are applied and costs are calculated for multiple-account jobs, see Understanding estimated usage costs.

7. Only the account that creates a job can access the job's details. This includes job-related details in the S3 bucket inventory.

8. Only the account that creates a job can access, suppress, or publish sensitive data findings that the job produces. Only an administrator can access, suppress, or publish sensitive data findings that automated sensitive data discovery produces.

9. If a sensitive data finding applies to an S3 object that a member account owns, the administrator might be able to retrieve samples of sensitive data reported by the finding. This depends on the source of the finding, and configuration settings and resources in the

administrator account and the member account. For more information, see Configuration options for retrieving sensitive data samples.

10. For an administrator to suspend Macie for their own account, the administrator must first disassociate their account from all member accounts.

11. For an administrator to disable Macie for their own account, the administrator must first disassociate their account from all member accounts, and delete the associations between their account and all of those accounts. The administrator for an organization in AWS Organizations can do this by working with the organization's management account to designate a different account as the administrator account.

    For a member of an AWS Organizations organization to disable Macie, the administrator must first disassociate the member's account from their administrator account. In an invitation-based organization, the member can disassociate their account from its administrator account, and then disable Macie.

12. The administrator for an organization in AWS Organizations can delete an association with a member account after they disassociate the account from their administrator account. The account continues to appear in the administrator's account inventory, but its status indicates that it's not a member account. In an invitation-based organization, an administrator and a member can delete an association with another account after they disassociate their account from the other account. The other account then stops appearing in their account inventory.

Managing multiple Macie accounts with AWS Organizations

If you use AWS Organizations to centrally manage multiple AWS accounts, you can integrate Amazon Macie with AWS Organizations, and then centrally manage Macie for accounts in your organization. With this configuration, a designated Macie administrator can enable and manage Macie for as many as 10,000 accounts. The administrator can also access Amazon Simple Storage Service (Amazon S3) inventory data and discover sensitive data in S3 buckets that the accounts own. For details about tasks that the administrator can perform, see Macie administrator and member account relationships.

AWS Organizations is a global account management service that enables AWS administrators to consolidate and centrally manage multiple AWS accounts. It provides account management and consolidated billing features that are designed to support budgetary, security, and compliance

needs. It's offered at no additional charge and it integrates with multiple AWS services, including Macie, AWS Security Hub, and Amazon GuardDuty. To learn more, see the AWS Organizations User Guide.

To integrate Macie with AWS Organizations, you start by designating an account as the delegated Macie administrator account for the organization. The Macie administrator then enables Macie for other accounts in the organization, adds those accounts as Macie member accounts, and configures Macie settings and resources for the accounts.

> **Tip**
>
> If you already associated a Macie administrator account with member accounts by using invitations, you can designate that account as the delegated Macie administrator account for your organization in AWS Organizations. If you do this, all currently associated member accounts remain members and you can take full advantage of the benefits of managing accounts by using AWS Organizations. For more information, see Transitioning from an invitation-based organization.

The topics in this section explain how to integrate Macie with AWS Organizations and how to administer and manage Macie for accounts in an organization.

Topics

Considerations for using Macie with AWS Organizations

Before you integrate Amazon Macie with AWS Organizations and configure your organization in Macie, consider the following requirements and recommendations. Also ensure that you understand the relationship between Macie administrator and member accounts.

Topics

- Designating a Macie administrator account
- Changing or removing the designation of a Macie administrator account
- Adding and removing Macie member accounts
- Transitioning from an invitation-based organization

Designating a Macie administrator account

While you determine which account should be the delegated Macie administrator account for your organization, keep the following in mind:

- An organization can have only one delegated Macie administrator account.
- An account can't be a Macie administrator and member account at the same time.
- Only the AWS Organizations management account for an organization can designate the delegated Macie administrator account for the organization. Only the management account can subsequently change or remove that designation.
- The AWS Organizations management account for an organization can also be the delegated Macie administrator account for the organization. However, we don't recommend this configuration based on AWS security best practices and the principle of least privilege. Users who have access to the management account for billing purposes are likely to be different from users who need access to Macie for information security purposes.

  If you prefer this configuration, you must enable Macie for the organization's management account in at least one AWS Region before you designate the account as the delegated Macie administrator account. Otherwise, the account won't be able to access and manage Macie settings and resources for member accounts.
- Unlike AWS Organizations, Macie is a Regional service. This means that the designation of a Macie administrator account is a Regional designation. It also means that associations between Macie administrator and member accounts are Regional. For example, if the management account designates a Macie administrator account in the US East (N. Virginia) Region, the Macie administrator can manage Macie for member accounts only in that Region.

  To centrally manage Macie accounts in multiple AWS Regions, the management account must sign in to each Region where the organization currently uses or will use Macie, and then designate the Macie administrator account in each of those Regions. The Macie administrator can then configure the organization in each of those Regions. For a list of Regions where Macie is currently available, see Amazon Macie endpoints and quotas in the *AWS General Reference*.

- An account can be associated with only one Macie administrator account at a time. If your organization uses Macie in multiple Regions, the designated Macie administrator account must be the same in all of those Regions. However, your organization's management account must designate the administrator account separately in each Region.
- An account can be the delegated Macie administrator account for only one organization at a time. If you manage multiple organizations in AWS Organizations, you must designate a different Macie administrator account for each organization. This is due to an AWS Organizations requirement—an account can be a member of only one organization at a time.

If the Macie administrator's AWS account is suspended, isolated, or closed, all associated Macie member accounts are automatically removed as Macie member accounts but Macie continues to be enabled for the accounts. If automated sensitive data discovery was enabled for one or more member accounts, it's disabled for the accounts. This also disables access to statistical data, inventory data, and other information that Macie produced and directly provided while performing automated discovery for the accounts. To restore access to this data, the following must occur within 30 days:

1. The Macie administrator's AWS account is restored.
2. The AWS Organizations management account designates the account as the Macie administrator account again.
3. The Macie administrator configures the organization and enables automated discovery for the appropriate accounts again.

After 30 days, Macie permanently deletes data that it previously produced and directly provided while performing automated discovery for the applicable accounts.

Changing or removing the designation of a Macie administrator account

Only the AWS Organizations management account for an organization can change or remove the designation of a delegated Macie administrator account for the organization.

If the management account changes or removes the designation:

- All associated member accounts are removed as Macie member accounts but Macie continues to be enabled for the accounts. The accounts become standalone Macie accounts. To pause or stop using Macie, a user of a member account must suspend (pause) or disable (stop) Macie for the account.

- Automated sensitive data discovery is disabled for each account that it was enabled for. This also disables access to statistical data, inventory data, and other information that Macie produced and directly provided while performing automated discovery for each account. To restore access to this data, the management account must designate the same Macie administrator account again within 30 days. In addition, the Macie administrator must configure the organization again and re-enable automated discovery for each account within 30 days. After 30 days, the data expires and Macie permanently deletes it.

Adding and removing Macie member accounts

As you add, remove, and otherwise manage member accounts for your organization, keep the following in mind:

- A Macie administrator account can be associated with no more than 10,000 Macie member accounts in each AWS Region. If your organization exceeds this quota, the Macie administrator won't be able to add member accounts until they remove the necessary number of existing member accounts in the Region. When an organization meets this quota, we notify the Macie administrator by creating an AWS Health event for their account. We also send email to the address that's associated with their account.

  If you're the Macie administrator for an organization, you can determine how many member accounts are currently associated with your account by using the **Accounts** page on the Amazon Macie console or the ListMembers operation of the Amazon Macie API. For more information, see Reviewing Macie accounts for an organization.
- An account can be associated with only one Macie administrator account at a time. This means that an account can't accept a Macie invitation from another account if it's already associated with the Macie administrator account for an organization in AWS Organizations.

  Similarly, if an account already accepted an invitation, the Macie administrator for an organization in AWS Organizations can't add the account as a Macie member account. The account must first disassociate from its current, invitation-based administrator account.
- To add the AWS Organizations management account as a Macie member account, a user of the management account must first enable Macie for the account. The Macie administrator isn't allowed to enable Macie for the management account.
- If the Macie administrator removes a Macie member account:

- Macie continues to be enabled for the account. The account becomes a standalone Macie account. To pause or stop using Macie, a user of the account must suspend (pause) or disable (stop) Macie for the account.
- Automated sensitive data discovery is disabled for the account, if it was enabled. This also disables access to statistical data, inventory data, and other information that Macie produced and directly provided while performing automated discovery for the account.

- A member account can't disassociate from its Macie administrator account. Only the Macie administrator can remove an account as a Macie member account.

Transitioning from an invitation-based organization

If you already associated a Macie administrator account with member accounts by using Macie membership invitations, we recommend that you designate that account as the delegated Macie administrator account for your organization in AWS Organizations. This simplifies the transition from an invitation-based organization.

If you do this, all currently associated member accounts continue to be members. If a member account is part of your organization in AWS Organizations, the account's association automatically changes from **By invitation** to **Via AWS Organizations** in Macie. If a member account isn't part of your organization in AWS Organizations, the account's association continues to be **By invitation**. In both cases, the accounts continue to be associated with the delegated Macie administrator account as member accounts. For sensitive data discovery, this also means that the accounts can continue to access statistical and other data that Macie produced and directly provided while performing automated sensitive data discovery for the accounts. In addition, if the Macie administrator configured sensitive data discovery jobs to analyze data for the accounts, subsequent job runs will continue to include resources that the accounts own.

We recommend this approach because an account can't be associated with more than one Macie administrator account at the same time. If you designate a different account as the Macie administrator account for your organization in AWS Organizations, the designated administrator won't be able to manage accounts that are already associated with another Macie administrator account by invitation. Each member account must first disassociate from its current, invitation-based administrator account. The Macie administrator for your organization in AWS Organizations can then add the account as a Macie member account and begin managing the account.

After you integrate Macie with AWS Organizations and you configure your organization in Macie, you can optionally designate a different Macie administrator account for the organization. You can

also continue to use invitations to associate and manage member accounts that aren't part of your organization in AWS Organizations.

Integrating and configuring an organization in Macie

To start using Amazon Macie with AWS Organizations, the AWS Organizations management account for the organization designates an account as the delegated Macie administrator account for the organization. This enables Macie as a trusted service in AWS Organizations. It also enables Macie in the current AWS Region for the designated administrator account, and it allows the designated administrator account to enable and manage Macie for other accounts in the organization in that Region. For information about how these permissions are granted, see Using AWS Organizations with other AWS services in the *AWS Organizations User Guide*.

The delegated Macie administrator then configures the organization in Macie, primarily by adding the organization's accounts as Macie member accounts in the Region. The administrator can then access certain Macie settings, data, and resources for those accounts in that Region. They can also perform automated sensitive data discovery and run sensitive data discovery jobs to detect sensitive data in Amazon Simple Storage Service (Amazon S3) buckets that the accounts own.

This topic explains how to designate a delegated Macie administrator for an organization and how to add the organization's accounts as Macie member accounts. Before you perform these tasks, ensure that you understand the relationship between Macie administrator and member accounts. It's also a good idea to review the considerations and recommendations for using Macie with AWS Organizations.

Tasks

- Step 1: Verify your permissions
- Step 2: Designate the delegated Macie administrator account for the organization
- Step 3: Automatically enable and add new organization accounts as Macie member accounts
- Step 4: Enable and add existing organization accounts as Macie member accounts

To integrate and configure the organization in multiple Regions, the AWS Organizations management account and the delegated Macie administrator repeat these steps in each additional Region.

Step 1: Verify your permissions

Before you designate the delegated Macie administrator account for your organization, verify that you (as a user of the AWS Organizations management account) are allowed to perform the following Macie action: `macie2:EnableOrganizationAdminAccount`. This action allows you to designate the delegated Macie administrator account for your organization by using Macie.

Also verify that you're allowed to perform the following AWS Organizations actions:

- `organizations:DescribeOrganization`
- `organizations:EnableAWSServiceAccess`
- `organizations:ListAWSServiceAccessForOrganization`
- `organizations:RegisterDelegatedAdministrator`

These actions allow you to: retrieve information about your organization; integrate Macie with AWS Organizations; retrieve information about which AWS services you've integrated with AWS Organizations; and, designate a delegated Macie administrator account for your organization.

To grant these permissions, include the following statement in an AWS Identity and Access Management (IAM) policy for your account:

```
{
   "Sid": "Grant permissions to designate a delegated Macie administrator",
   "Effect": "Allow",
   "Action": [
      "macie2:EnableOrganizationAdminAccount",
      "organizations:DescribeOrganization",
      "organizations:EnableAWSServiceAccess",
      "organizations:ListAWSServiceAccessForOrganization",
      "organizations:RegisterDelegatedAdministrator"
   ],
   "Resource": "*"
}
```

If you want to designate your AWS Organizations management account as the delegated Macie administrator account for the organization, your account also needs permission to perform the following IAM action: `CreateServiceLinkedRole`. This action allows you to enable Macie for the management account. However, based on AWS security best practices and the principle of least privilege, we don't recommend that you do this.

If you decide to grant this permission, add the following statement to the IAM policy for your AWS Organizations management account:

```
{
   "Sid": "Grant permissions to enable Macie",
   "Effect": "Allow",
   "Action": [
      "iam:CreateServiceLinkedRole"
   ],
   "Resource": "arn:aws:iam::111122223333:role/aws-service-role/macie.amazonaws.com/
AWSServiceRoleForAmazonMacie",
   "Condition": {
      "StringLike": {
         "iam:AWSServiceName": "macie.amazonaws.com"
      }
   }
}
```

In the statement, replace *111122223333* with the account ID for the management account.

If you want to administer Macie in an opt-in AWS Region (Region that's disabled by default), also update the value for the Macie service principal in the `Resource` element and the `iam:AWSServiceName` condition. The value must specify the Region code for the Region. For example, to administer Macie in the Middle East (Bahrain) Region, which has the Region code *me-south-1*, do the following:

- In the `Resource` element, replace

  `arn:aws:iam::111122223333:role/aws-service-role/macie.amazonaws.com/AWSServiceRoleForAmazonMacie`

  with

  `arn:aws:iam::111122223333:role/aws-service-role/macie.me-south-1.amazonaws.com/AWSServiceRoleForAmazonMacie`

  Where *111122223333* specifies the account ID for the management account and *me-south-1* specifies the Region code for the Region.
- In the `iam:AWSServiceName` condition, replace `macie.amazonaws.com` with `macie.me-south-1.amazonaws.com`, where *me-south-1* specifies the Region code for the Region.

For a list of Regions where Macie is currently available and the Region code for each one, see Amazon Macie endpoints and quotas in the *AWS General Reference*. To determine whether a Region is an opt-in Region, see Enable or disable AWS Regions in your account in the *AWS Account Management User Guide*.

Step 2: Designate the delegated Macie administrator account for the organization

After you verify your permissions, you (as a user of the AWS Organizations management account) can designate the delegated Macie administrator account for your organization.

To designate the delegated Macie administrator account for an organization

To designate the delegated Macie administrator account for your organization, you can use the Amazon Macie console or the Amazon Macie API. Only a user of the AWS Organizations management account can perform this task.

Console

Follow these steps to designate the delegated Macie administrator account by using the Amazon Macie console.

To designate the delegated Macie administrator account

1. Sign in to the AWS Management Console using your AWS Organizations management account.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to designate the delegated Macie administrator account for your organization.
3. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
4. Do one of the following, depending on whether Macie is enabled for your management account in the current Region:
   - If Macie isn't enabled, choose **Get started** on the welcome page.
   - If Macie is enabled, choose **Settings** in the navigation pane.
5. Under **Delegated administrator**, enter the 12-digit account ID for the AWS account that you want to designate as the Macie administrator account.
6. Choose **Delegate**.

Repeat the preceding steps in each additional Region in which you want to integrate your organization with Macie. You must designate the same Macie administrator account in each of those Regions.

API

To designate the delegated Macie administrator account programmatically, use the EnableOrganizationAdminAccount operation of the Amazon Macie API. To designate the account in multiple Regions, submit the designation for each Region in which you want to integrate your organization with Macie. You must designate the same Macie administrator account in each of those Regions.

When you submit the designation, use the required `adminAccountId` parameter to specify the 12-digit account ID for the AWS account to designate as the Macie administrator account for the organization. Also ensure that you specify the Region that the designation applies to.

To designate the Macie administrator account by using the AWS Command Line Interface (AWS CLI), run the enable-organization-admin-account command. For the `admin-account-id` parameter, specify the 12-digit account ID for the AWS account to designate. Use the `region` parameter to specify the Region that the designation applies to. For example:

```
C:\> aws macie2 enable-organization-admin-account --region us-east-1 --admin-
account-id 111122223333
```

Where *us-east-1* is the Region that the designation applies to (the US East (N. Virginia) Region) and *111122223333* is the account ID for the account to designate.

After you designate the Macie administrator account for your organization, the Macie administrator can begin configuring the organization in Macie.

Step 3: Automatically enable and add new organization accounts as Macie member accounts

By default, Macie isn't automatically enabled for new accounts when the accounts are added to your organization in AWS Organizations. In addition, the accounts aren't automatically added as Macie member accounts. The accounts appear in the Macie administrator's account inventory. However, Macie isn't necessarily enabled for the accounts and the Macie administrator can't necessarily access Macie settings, data, and resources for the accounts.

If you're the delegated Macie administrator for the organization, you can change this configuration setting. You can turn on automatic enablement for your organization. If you do this, Macie is automatically enabled for new accounts when the accounts are added to your organization in AWS Organizations. In addition, the accounts are automatically associated with your Macie administrator account as member accounts. Turning on this setting doesn't affect existing accounts in your organization. To enable and manage Macie for existing accounts, you must manually add the accounts as Macie member accounts. The next step explains how to do this.

> **Note**
>
> If you turn on automatic enablement, note the following exception. If a new account is already associated with a different Macie administrator account, Macie doesn't automatically add the account as a member account in your organization. The account must disassociate from its current Macie administrator account before it can be part of your organization in Macie. You can then manually add the account. To identify accounts where this is the case, you can review the account inventory for your organization.

To automatically enable and add new organization accounts as Macie member accounts

To automatically enable and add new accounts as Macie member accounts, you can use the Amazon Macie console or the Amazon Macie API. Only the delegated Macie administrator for the organization can perform this task.

Console

To perform this task by using the console, you must be allowed to perform the following AWS Organizations action: `organizations:ListAccounts`. This action allows you to retrieve and display information about the accounts in your organization. If you have these permissions, follow these steps to automatically enable and add new organization accounts as Macie member accounts.

To automatically enable and add new organization accounts

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to automatically enable and add new accounts as Macie member accounts.
3. In the navigation pane, choose **Accounts**.

4. On the **Accounts** page, in the **New accounts** section, choose **Edit**.
5. In the **Edit settings for new accounts** dialog box, select **Enable Macie**.

   To also enable automated sensitive data discovery automatically for new member accounts, select **Enable automated sensitive data discovery**. If you enable this feature for an account, Macie continually selects sample objects from the account's S3 buckets and analyzes the objects to determine whether they contain sensitive data. For more information, see Performing automated sensitive data discovery.
6. Choose **Save**.

Repeat the preceding steps in each additional Region in which you want to configure your organization in Macie.

To subsequently change these settings, repeat the preceding steps and clear the checkbox for each setting.

API

To automatically enable and add new Macie member accounts programmatically, use the UpdateOrganizationConfiguration operation of the Amazon Macie API. When you submit your request, set the value for the `autoEnable` parameter to `true`. (The default value is `false`.) Also ensure that you specify the Region that your request applies to. To automatically enable and add new accounts in additional Regions, submit the request for each additional Region.

If you use the AWS CLI to submit the request, run the update-organization-configuration command and specify the `auto-enable` parameter to enable and add new accounts automatically. For example:

```
$ aws macie2 update-organization-configuration --region us-east-1 --auto-enable
```

Where *us-east-1* is the Region in which to automatically enable and add new accounts, the US East (N. Virginia) Region.

To subsequently change this setting and stop enabling and adding new accounts automatically, run the same command again and use the `no-auto-enable` parameter, instead of the `auto-enable` parameter, in each applicable Region.

You can also enable automated sensitive data discovery automatically for new member accounts. If you enable this feature for an account, Macie continually selects sample

objects from the account's S3 buckets and analyzes the objects to determine whether they contain sensitive data. For more information, see Performing automated sensitive data discovery. To enable this feature automatically for member accounts, use the UpdateAutomatedDiscoveryConfiguration operation or, if you're using the AWS CLI, run the update-automated-discovery-configuration command.

Step 4: Enable and add existing organization accounts as Macie member accounts

When you integrate Macie with AWS Organizations, Macie isn't automatically enabled for all the existing accounts in your organization. In addition, the accounts aren't automatically associated with the delegated Macie administrator account as Macie member accounts. Therefore, the final step of integrating and configuring your organization in Macie is to add existing organization accounts as Macie member accounts. When you add an existing account as a Macie member account, Macie is automatically enabled for the account and you (as the delegated Macie administrator) gain access to certain Macie settings, data, and resources for the account.

Note that you can't add an account that's currently associated with another Macie administrator account. To add the account, work with the account owner to first disassociate the account from its current administrator account. In addition, you can't add an existing account if Macie is currently suspended for the account. The account owner must first re-enable Macie for the account. Finally, if you want to add the AWS Organizations management account as a member account, a user of that account must first enable Macie for the account.

To enable and add existing organization accounts as Macie member accounts

To enable and add existing organization accounts as Macie member accounts, you can use the Amazon Macie console or the Amazon Macie API. Only the delegated Macie administrator for the organization can perform this task.

Console

To perform this task by using the console, you must be allowed to perform the following AWS Organizations action: `organizations:ListAccounts`. This action allows you to retrieve and display information about the accounts in your organization. If you have these permissions, follow these steps to enable and add existing accounts as Macie member accounts.

To enable and add existing organization accounts

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.

2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to enable and add existing accounts as Macie member accounts.
3. In the navigation pane, choose **Accounts**. The **Accounts** page opens and displays a table of the accounts that are associated with your Macie account.

   If an account is part of your organization in AWS Organizations, its **Type** is **Via AWS Organizations**. If an account is already a Macie member account, its **Status** is **Enabled** or **Paused (suspended)**.
4. In the **Existing accounts** table, select the checkbox for each account that you want to add as a Macie member account.
5. On the **Actions** menu, choose **Add member**.
6. Confirm that you want to add the selected accounts as member accounts.

After you confirm the addition of the selected accounts, the status of the accounts changes to **Enabling in process** and then **Enabled**. After you add a member account, you can also enable automated sensitive data discovery for the account: in the **Existing accounts** table, select the checkbox for each account to enable it for, and then choose **Enable automated sensitive data discovery** on the **Actions** menu. If you enable this feature for an account, Macie continually selects sample objects from the account's S3 buckets and analyzes the objects to determine whether they contain sensitive data. For more information, see Performing automated sensitive data discovery.

Repeat the preceding steps in each additional Region in which you want to configure your organization in Macie.

API

To programmatically enable and add one or more existing accounts as Macie member accounts, use the CreateMember operation of the Amazon Macie API. When you submit your request, use the supported parameters to specify the 12-digit account ID and email address of each AWS account to enable and add. Also specify the Region that the request applies to. To enable and add existing accounts in additional Regions, submit the request for each additional Region.

To retrieve the account ID and email address of an AWS account to enable and add, you can optionally use the ListMembers operation of the Amazon Macie API. This operation provides details about the accounts that are associated with your Macie account, including accounts that aren't Macie member accounts. If the value for the `relationshipStatus` property of an account isn't `Enabled` or `Paused`, the account isn't a Macie member account.

To enable and add one or more existing accounts by using the AWS CLI, run the create-member command. Use the `region` parameter to specify the Region in which to enable and add the accounts. Use the `account` parameters to specify the account ID and email address for each AWS account to add. For example:

```
C:\> aws macie2 create-member --region us-east-1 --account={\"accountId\":
\"123456789012\",\"email\":\"janedoe@example.com\"}
```

Where *us-east-1* is the Region in which to enable and add the account as a Macie member account (the US East (N. Virginia) Region), and the `account` parameters specify the account ID (*123456789012*) and email address (*janedoe@example.com*) for the account.

If your request succeeds, the status (`relationshipStatus`) of the specified account changes to `Enabled` in your account inventory.

To also enable automated sensitive data discovery for one or more of the accounts, use the BatchUpdateAutomatedDiscoveryAccounts operation or, if you're using the AWS CLI, run the batch-update-automated-discovery-accounts command. If you enable this feature for an account, Macie continually selects sample objects from the account's S3 buckets and analyzes the objects to determine whether they contain sensitive data. For more information, see Performing automated sensitive data discovery.

Reviewing Macie accounts for an organization

After an AWS Organizations organization is integrated and configured in Amazon Macie, the delegated Macie administrator can access an inventory of the organization's accounts in Macie. As the Macie administrator for an organization, you can use this inventory to review statistics and details for your organization's Macie accounts in an AWS Region. You can also use it to perform certain management tasks for the accounts.

To review the Macie accounts for an organization

To review the accounts for your organization, you can use the Amazon Macie console or the Amazon Macie API. If you prefer to use the console, you must be allowed to perform the following AWS Organizations action: `organizations:ListAccounts`. This action allows you to retrieve and display information about accounts that are part of your organization in AWS Organizations.

Console

Follow these steps to review your organization's Macie accounts by using the Amazon Macie console.

To review your organization's accounts

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to review your organization's accounts.
3. In the navigation pane, choose **Accounts**.

The **Accounts** page opens and displays aggregated statistics and a table of the accounts that are associated with your Macie account in the current AWS Region.

At the top of the **Accounts** page, you'll find the following aggregated statistics.

Via AWS Organizations

Active reports the total number of accounts that are associated with your account through AWS Organizations and are currently Macie member accounts in your organization. Macie is enabled for these accounts and you're the Macie administrator of the accounts.

All reports the total number of accounts that are associated with your account through AWS Organizations. This includes accounts that aren't currently Macie member accounts. It also includes member accounts that Macie is currently suspended for.

By invitation

Active reports the total number of accounts that are associated with your account by Macie invitation and are currently Macie member accounts in your organization. These accounts aren't associated with your account through AWS Organizations. Macie is enabled for the accounts and you're the Macie administrator of the accounts because they accepted a Macie membership invitation from you.

All reports the total number of accounts that are associated with your account by Macie invitation, including accounts that haven't responded to an invitation from you.

Active/All

Active reports the total number of accounts that Macie is currently enabled for in your organization, including your own account. You're the Macie administrator of these accounts through AWS Organizations or by Macie invitation.

All reports the total number of accounts that are associated with your account, through AWS Organizations or by Macie invitation, plus your own account. This includes accounts that are part of your organization in AWS Organizations and aren't currently Macie member accounts. It also includes any accounts that haven't responded to a Macie membership invitation from you.

In the table, you'll find details about each account in the current Region. The table includes all the accounts that are associated with your Macie account through AWS Organizations or by Macie invitation.

Account ID

The account ID and email address for the AWS account.

Name

The account name for the AWS account. This value is typically **N/A** for your own account, and any accounts that are associated with your account by Macie invitation.

Type

How the account is associated with your account, through AWS Organizations or by Macie invitation. For your own account, this value is **Current account**.

Status

The status of the relationship between your account and the account. For an account in an AWS Organizations organization (**Type** is **Via AWS Organizations**), possible values are:

- **Account suspended** – The AWS account is suspended.
- **Enabled** – The account is a Macie member account. Macie is enabled for the account and you're the Macie administrator of the account.
- **Enabling in process** – Macie is processing a request to enable and add the account as a Macie member account.
- **Not a member** – The account is part of your organization in AWS Organizations but it isn't a Macie member account.

- **Paused (suspended)** – The account is a Macie member account but Macie is currently suspended for the account.
- **Region disabled** – The account is part of your organization in AWS Organizations but the current Region is disabled for the AWS account.
- **Removed (disassociated)** – The account was previously a Macie member account but was subsequently removed as a member account. You disassociated the account from your Macie administrator account. Macie continues to be enabled for the account.

Last status update

When you or the associated account most recently performed an action that affected the relationship between your accounts.

Automated sensitive data discovery

Whether automated sensitive data discovery is currently enabled or disabled for the account.

To sort the table by a specific field, choose the column heading for the field. To change the sort order, choose the column heading again. To filter the table, place your cursor in the filter box, and then add a filter condition for a field. To further refine the results, add filter conditions for additional fields.

API

To review your organization's accounts programmatically, use the ListMembers operation of the Amazon Macie API and specify the Region that your request applies to. To review the accounts in additional Regions, submit your request in each additional Region.

When you submit your request, use the `onlyAssociated` parameter to specify which accounts to include in the response. By default, Macie returns details about only those accounts that are Macie member accounts in the specified Region through AWS Organizations or by Macie invitation. To retrieve these details for all the accounts that are associated with your Macie account, including accounts that aren't member accounts, include the `onlyAssociated` parameter in your request and set the parameter's value to `false`.

To review your organization's accounts by using the AWS Command Line Interface (AWS CLI), run the list-members command. For the `only-associated` parameter, specify whether to include all associated accounts or only Macie member accounts. To include only member accounts, omit this parameter or set the parameter's value to `true`. To include all accounts, set this value to `false`. For example:

```
C:\> aws macie2 list-members --region us-east-1 --only-associated false
```

Where *us-east-1* is the Region that the request applies to, the US East (N. Virginia) Region.

If your request succeeds, Macie returns a `members` array. The array contains a `member` object for each account that meets the criteria specified in the request. In that object, the `relationshipStatus` field indicates the current status of the relationship between your account and the other account in the specified Region. For an account in an AWS Organizations organization, possible values are:

- `AccountSuspended` – The AWS account is suspended.
- `Created` – Macie is processing a request to enable and add the account as a Macie member account.
- `Enabled` – The account is a Macie member account. Macie is enabled for the account and you're the Macie administrator of the account.
- `Paused` – The account is a Macie member account but Macie is currently suspended (paused) for the account.
- `RegionDisabled` – The account is part of your organization in AWS Organizations but the current Region is disabled for the AWS account.
- `Removed` – The account was previously a Macie member account but was subsequently removed as a member account. You disassociated the account from your Macie administrator account. Macie continues to be enabled for the account.

For information about other fields in the `member` object, see Members in the *Amazon Macie API Reference*.

Managing Macie member accounts for an organization

After an AWS Organizations organization is integrated and configured in Amazon Macie, the organization's delegated Macie administrator can access certain Macie settings, data, and resources for member accounts. As the Macie administrator for an organization, you can use Macie to centrally perform certain account management and administration tasks for the accounts. For example, you can:

- Add and remove accounts as Macie member accounts.

- Manage the status of Macie for individual accounts, such as enable or suspend Macie for an account.
- Monitor Macie quotas and estimated usage costs for individual accounts and the organization overall.

You can also review Amazon Simple Storage Service (Amazon S3) inventory data and policy findings for Macie member accounts. And you can discover sensitive data in S3 buckets that the accounts own. For a detailed list of tasks that you can perform, see Macie administrator and member account relationships.

By default, Macie gives you visibility into relevant data and resources for all the Macie member accounts in your organization. You can also drill down to review data and resources for individual accounts. For example, if you use the Summary dashboard to assess your organization's Amazon S3 security posture, you can filter the data by account. Similarly, if you monitor estimated usage costs, you can access breakdowns of estimated costs for individual member accounts.

In addition to tasks that are common to administrator and member accounts, you can perform various administrative tasks for your organization.

Tasks

- Adding Macie member accounts to an organization
- Suspending Macie for member accounts in an organization
- Removing Macie member accounts from an organization

As the Macie administrator for an organization, you can perform these tasks by using the Amazon Macie console or the Amazon Macie API. If you prefer to use the console, you must be allowed to perform the following AWS Organizations action: `organizations:ListAccounts`. This action allows you to retrieve and display information about accounts that are part of your organization in AWS Organizations.

Adding Macie member accounts to an organization

In some cases, you might need to manually add an account as an Amazon Macie member account. This is the case for accounts that you previously removed (disassociated) as member accounts. This is also the case if you didn't configure Macie to automatically enable and add new member accounts when accounts are added to your organization in AWS Organizations.

When you add an account as a Macie member account:

- Macie is enabled for the account in the current AWS Region, if it isn't already enabled in the Region.
- The account is associated with your Macie administrator account as a member account in the Region. The member account doesn't receive an invitation or other notification that you established this relationship between your accounts.
- Automated sensitive data discovery might be enabled for the account in the Region. This depends on configuration settings that you specified for the organization. For more information, see Configuring automated sensitive data discovery.

Note that you can't add an account that's already associated with another Macie administrator account. The account must first disassociate from its current administrator account. In addition, you can't add the AWS Organizations management account as a member account unless Macie is already enabled for the account. To learn about additional requirements, see Considerations for using Macie with AWS Organizations.

To add a Macie member account to an organization

To add one or more Macie member accounts to your organization, you can use the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to add one or more Macie member accounts by using the Amazon Macie console.

To add a Macie member account

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to add a member account.
3. In the navigation pane, choose **Accounts**. The **Accounts** page opens and displays a table of the accounts that are associated with your account.
4. (Optional) To more easily identify accounts that are part of your organization in AWS Organizations and aren't Macie member accounts, use the filter box above the **Existing accounts** table to add the following filter conditions:

- **Type = Organization**
- **Status = Not a Member**

To also display accounts that you previously removed and might want to add as member accounts, also add a **Status = Removed** filter condition.

5. In the **Existing accounts** table, select the checkbox for each account that you want to add as a member account.
6. On the **Actions** menu, choose **Add member**.
7. Confirm that you want to add the selected accounts as member accounts.

After you confirm your selections, the status of the selected accounts changes to **Enabling in process**, and then **Enabled** in your account inventory.

To add a member account in additional Regions, repeat the preceding steps in each additional Region.

API

To add one or more Macie member accounts programmatically, use the CreateMember operation of the Amazon Macie API.

When you submit your request, use the supported parameters to specify the 12-digit account ID and email address for each AWS account that you want to add. Also specify the Region that the request applies to. To add an account in additional Regions, submit your request in each additional Region.

To retrieve the account ID and email address of an account to add, you can correlate the output of the ListAccounts operation of the AWS Organizations API and the ListMembers operation of the Amazon Macie API. For the **ListMembers** operation of the Macie API, include the `onlyAssociated` parameter in your request and set the parameter's value to `false`. If the operation succeeds, Macie returns a `members` array that provides details about all the accounts that are associated with your Macie administrator account in the specified Region, including accounts that aren't currently member accounts. Note the following in the array:

- If the value for the `relationshipStatus` property of an account isn't `Enabled` or `Paused`, the account is associated with your account but it isn't a Macie member account.

- If an account isn't included in the array but is included in the output of the **ListAccounts** operation of the AWS Organizations API, the account is part of your organization in AWS Organizations but it isn't associated with your account and, therefore, isn't a Macie member account.

To add a member account by using the AWS Command Line Interface (AWS CLI), run the create-member command. Use the `region` parameter to specify the Region in which to add the account. Use the `account` parameters to specify the account ID and email address for each account to add. For example:

```
C:\> aws macie2 create-member --region us-east-1 --account={\"accountId\":
\"123456789012\",\"email\":\"janedoe@example.com\"}
```

Where *us-east-1* is the Region in which to add the account as a member account (the US East (N. Virginia) Region), and the `account` parameters specify the account ID (*123456789012*) and email address (*janedoe@example.com*) for the account.

If your request succeeds, the status (`relationshipStatus`) of the specified account changes to `Enabled` in your account inventory.

Suspending Macie for member accounts in an organization

As the Amazon Macie administrator for an organization in AWS Organizations, you can suspend Macie for a member account in your organization. If you do this, you can also re-enable Macie for the account at a later time.

When you suspend Macie for a member account:

- Macie loses access to and stops providing metadata about the account's Amazon S3 data in the current AWS Region.
- Macie stops performing all activities for the account in the Region. This includes monitoring S3 buckets for security and access control, performing automated sensitive data discovery, and running sensitive data discovery jobs that are currently in progress.
- Macie cancels all sensitive data discovery jobs that were created by the account in the Region. A job can't be resumed or restarted after it's cancelled. If you created jobs to analyze data that the member account owns, Macie doesn't cancel your jobs. Instead, the jobs skip resources that are owned by the account.

While it's suspended, Macie retains the session identifier, settings, and resources that it stores or maintains for the account in the applicable Region. Macie also retains certain data for the account in the Region. For example, the account's findings remain intact and aren't affected for up to 90 days. If automated sensitive data discovery was enabled for the account, existing results also remain intact and aren't affected for up to 30 days. Your organization doesn't incur Macie charges for the account in that Region while Macie is suspended for the account in the Region.

To suspend Macie for a member account in an organization

To suspend Macie for a member account in an organization, you can use the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to suspend Macie for a member account by using the Amazon Macie console.

To suspend Macie for a member account

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to suspend Macie for a member account.
3. In the navigation pane, choose **Accounts**. The **Accounts** page opens and displays a table of the accounts that are associated with your account.
4. In the **Existing accounts** table, select the checkbox for the account to suspend Macie for.
5. On the **Actions** menu, choose **Suspend Macie**.
6. Confirm that you want to suspend Macie for the account.

After you confirm the suspension, the status of the account changes to **Paused (suspended)** in your account inventory. To suspend Macie for the account in additional Regions, repeat the preceding steps in each additional Region.

To later re-enable Macie for the account, return to the **Accounts** page on the console. Select the checkbox for the account, and then choose **Enable Macie** on the **Actions** menu. To re-enable Macie for the account in additional Regions, repeat these steps in each additional Region.

API

To suspend Macie for a member account programmatically, use the UpdateMemberSession operation of the Amazon Macie API. You can also use this operation to later re-enable Macie for the account.

When you submit your request, use the `id` parameter to specify the 12-digit account ID for the AWS account that you want to suspend Macie for. For the `status` parameter, specify `PAUSED`. Also specify the Region that the request applies to. To suspend Macie for the account in additional Regions, submit your request in each additional Region.

To retrieve the account ID for the account, you can use the ListMembers operation of the Amazon Macie API. If you do this, consider filtering the results by including the `onlyAssociated` parameter in your request. If you set this parameter's value to `true`, Macie returns a `members` array that provides details about only those accounts that are currently member accounts.

To suspend Macie for a member account by using the AWS CLI, run the update-member-session command. Use the `region` parameter to specify the Region in which to suspend Macie for the account. Use the `id` parameter to specify the account ID for the account. For the `status` parameter, specify `PAUSED`. For example:

```
C:\> aws macie2 update-member-session --region us-east-1 --id 123456789012 --status
 PAUSED
```

Where *`us-east-1`* is the Region in which to suspend Macie (the US East (N. Virginia) Region), *`123456789012`* is the account ID for the account to suspend Macie for, and `PAUSED` is the new status of Macie for the account.

If your request succeeds, Macie returns an empty response and the status of the specified account changes to `Paused` in your account inventory. To later re-enable Macie for the account, run the **update-member-session** command again and specify `ENABLED` for the `status` parameter.

Removing Macie member accounts from an organization

If you want to stop accessing Amazon Macie settings, data, and resources for a member account, you can remove the account as a Macie member account. You do this by disassociating the account

from your Macie administrator account. Note that only you can do this for a member account. An AWS Organizations member account can't disassociate from its Macie administrator account.

When you remove a Macie member account, Macie remains enabled for the account in the current AWS Region. However, the account is disassociated from your Macie administrator account and it becomes a standalone Macie account. This means that you lose access to all Macie settings, data, and resources for the account, including metadata and policy findings for the account's Amazon S3 data. This also means that you can no longer use Macie to discover sensitive data in S3 buckets that the account owns. If you already created sensitive data discovery jobs to do this, the jobs skip buckets that the account owns. If you enabled automated sensitive data discovery for the account, both you and the member account lose access to statistical data, inventory data, and other information that Macie produced and directly provided while performing automated discovery for the account.

After you remove a Macie member account, the account continues to appear in your account inventory. Macie doesn't notify the account's owner that you removed the account. Therefore, consider contacting the account owner to ensure that they begin managing settings and resources for their account.

You can add the account to your organization again at a later time. If you do this and you enable automated sensitive data discovery for the account again within 30 days, you also regain access to data and information that Macie previously produced and directly provided while performing automated discovery for the account. In addition, subsequent runs of your existing jobs start including the account's S3 buckets again.

To remove a Macie member account from an organization

To remove a Macie member account from your organization, you can use the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to remove a Macie member account by using the Amazon Macie console.

To remove a Macie member account

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to remove a member account.

3. In the navigation pane, choose **Accounts**. The **Accounts** page opens and displays a table of the accounts that are associated with your account.
4. In the **Existing accounts** table, select the checkbox for the account that you want to remove as a member account.
5. On the **Actions** menu, choose **Disassociate account**.
6. Confirm that you want to remove the selected account as a member account.

After you confirm your selection, the status of the account changes to **Removed (disassociated)** in your account inventory.

To remove the member account in additional Regions, repeat the preceding steps in each additional Region.

API

To remove a Macie member account programmatically, use the DisassociateMember operation of the Amazon Macie API.

When you submit your request, use the `id` parameter to specify the 12-digit AWS account ID for the member account to remove. Also specify the Region that the request applies to. To remove the account in additional Regions, submit your request in each additional Region.

To retrieve the account ID for the member account to remove, you can use the ListMembers operation of the Amazon Macie API. If you do this, consider filtering the results by including the `onlyAssociated` parameter in your request. If you set this parameter's value to `true`, Macie returns a `members` array that provides details about only those accounts that are currently Macie member accounts.

To remove a Macie member account by using the AWS CLI, run the disassociate-member command. Use the `region` parameter to specify the Region in which to remove the account. Use the `id` parameter to specify the account ID for the member account to remove. For example:

```
C:\> aws macie2 disassociate-member --region us-east-1 --id 123456789012
```

Where *us-east-1* is the Region in which to remove the account (the US East (N. Virginia) Region) and *123456789012* is the account ID for the account to remove.

If your request succeeds, Macie returns an empty response and the status of the specified account changes to `Removed` in your account inventory.

Changing the Macie administrator account for an organization

After an AWS Organizations organization is integrated and configured in Amazon Macie, the AWS Organizations management account can designate a different account as the delegated Macie administrator account for the organization. The new Macie administrator can then configure the organization in Macie again.

As a user of the AWS Organizations management account for an organization, verify that you meet the following permissions requirements before you designate a different Macie administrator account for your organization:

- You must have the same permissions that were required to initially designate a Macie administrator account for your organization. You must also be allowed to perform the following AWS Organizations action: `organizations:DeregisterDelegatedAdministrator`. This additional action allows you to remove the current designation.
- If your account is currently a Macie member account, the current Macie administrator must remove your account as a Macie member account. Otherwise, you won't be allowed to access Macie operations for designating a different administrator account. After you designate a new administrator account, the new Macie administrator can add your account as a Macie member account again.

If your organization uses Macie in multiple AWS Regions, also ensure that you change the designation in each Region in which your organization uses Macie. The delegated Macie administrator account must be the same in all of those Regions. If you manage multiple organizations in AWS Organizations, also note that an account can be the delegated Macie administrator account for only one organization at a time. To learn about additional requirements, see Considerations for using Macie with AWS Organizations.

> **Note**
>
> When you designate a different Macie administrator account for your organization, you also disable access to existing statistical data, inventory data, and other information that Macie produced and directly provided while performing automated sensitive data discovery for accounts in the organization. The new Macie administrator can't access the existing data. If

you change the designation and the new Macie administrator enables automated discovery for the accounts, Macie generates and maintains new data when it performs automated discovery for the accounts.

To change the designation of a Macie administrator account

To designate a different Macie administrator account for your organization, you can use the Amazon Macie console or a combination of the Amazon Macie and AWS Organizations APIs. Only a user of the AWS Organizations management account can change the designation for their organization.

Console

Follow these steps to change the designation by using the Amazon Macie console.

To change the designation

1. Sign in to the AWS Management Console by using your AWS Organizations management account.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to change the designation.
3. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
4. Do one of the following, depending on whether Macie is enabled for your management account in the current Region:
   - If Macie isn't enabled, choose **Get started** on the welcome page.
   - If Macie is enabled, choose **Settings** in the navigation pane.
5. Under **Delegated administrator**, choose **Remove**. To change the designation, you must first remove the current designation.
6. Confirm that you want to remove the current designation.
7. Under **Delegated administrator**, enter the 12-digit account ID for the AWS account to designate as the new Macie administrator account for the organization.
8. Choose **Delegate**.

Repeat the preceding steps in each additional Region in which you integrated Macie with AWS Organizations.

API

To change the designation programmatically, you use two operations of the Amazon Macie API and one operation of the AWS Organizations API. This is because you have to remove the current designation in both Macie and AWS Organizations before you submit the new designation.

To remove the current designation:

1. Use the DisableOrganizationAdminAccount operation of the Macie API. For the required `adminAccountId` parameter, specify the 12-digit account ID for the AWS account that's currently designated as the Macie administrator account for the organization.
2. Use the DeregisterDelegatedAdministrator operation of the AWS Organizations API. For the `AccountId` parameter, specify the 12-digit account ID for the account that's currently designated as the Macie administrator account for the organization. This value should match the account ID that you specified in the preceding Macie request. For the `ServicePrincipal` parameter, specify the Macie service principal (`macie.amazonaws.com`).

After you remove the current designation, submit the new designation by using the EnableOrganizationAdminAccount operation of the Macie API. For the required `adminAccountId` parameter, specify the 12-digit account ID for the AWS account to designate as the new Macie administrator account for the organization.

To change the designation by using the AWS Command Line Interface (AWS CLI), run the disable-organization-admin-account command of the Macie API and the deregister-delegated-administrator command of the AWS Organizations API. These commands remove the current designation in Macie and AWS Organizations, respectively. For the `admin-account-id` and `account-id` parameters, specify the 12-digit account ID for the AWS account to remove as the current Macie administrator account. Use the `region` parameter to specify the Region that the removal applies to. For example:

```
C:\> aws macie2 disable-organization-admin-account --region us-east-1 --admin-
account-id 111122223333 && aws organizations deregister-delegated-administrator --
region us-east-1 --account-id 111122223333 --service-principal macie.amazonaws.com
```

Where:

- *us-east-1* is the Region that the removal applies to, the US East (N. Virginia) Region.
- *111122223333* is the account ID for the account to remove as the Macie administrator account.
- `macie.amazonaws.com` is the Macie service principal.

After you remove the current designation, submit the new designation by running the enable-organization-admin-account command of the Macie API. For the `admin-account-id` parameter, specify the 12-digit account ID for the AWS account to designate as the new Macie administrator account for the organization. Use the `region` parameter to specify the Region that the designation applies to. For example:

```
C:\> aws macie2 enable-organization-admin-account --region us-east-1 --admin-
account-id 444455556666
```

Where *us-east-1* is the Region that the designation applies to (the US East (N. Virginia) Region) and *444455556666* is the account ID for the account to designate as the new Macie administrator account.

Disabling Macie integration with AWS Organizations

After an AWS Organizations organization is integrated with Amazon Macie, the AWS Organizations management account can subsequently disable the integration. As a user of the AWS Organizations management account, you can do this by disabling trusted service access for Macie in AWS Organizations.

When you disable trusted service access for Macie, the following occurs:

- Macie loses its status as a trusted service in AWS Organizations.
- The organization's Macie administrator account loses access to all Macie settings, data, and resources for all Macie member accounts in all AWS Regions.
- All Macie member accounts become standalone Macie accounts. If Macie was enabled for a member account in one or more Regions, Macie continues to be enabled for the account in those Regions. However, the account is no longer associated with a Macie administrator account in any Region. In addition, the account loses access to statistical data, inventory data, and other information that Macie produced and directly provided while performing automated sensitive data discovery for the account.

For additional information about the results of disabling trusted service access, see Using AWS Organizations with other AWS services in the *AWS Organizations User Guide*.

To disable trusted service access for Macie

To disable trusted service access, you can use the AWS Organizations console or the AWS Organizations API. Only a user of the AWS Organizations management account can disable trusted service access for Macie. For details about the permissions that you need, see Permissions required to disable trusted access in the *AWS Organizations User Guide*.

Before you disable trusted service access, optionally work with the delegated Macie administrator for your organization to suspend or disable Macie for member accounts and to clean up Macie resources for the accounts.

Console

To disable trusted service access by using the AWS Organizations console, follow these steps.

To disable trusted service access

1. Sign in to the AWS Management Console using your AWS Organizations management account.
2. Open the AWS Organizations console at https://console.aws.amazon.com/organizations/.
3. In the navigation pane, choose **Services**.
4. Under **Integrated services**, choose **Amazon Macie**.
5. Choose **Disable trusted access**.
6. Confirm that you want to disable trusted access.

API

To disable trusted service access programmatically, use the DisableAWSServiceAccess operation of the AWS Organizations API. For the `ServicePrincipal` parameter, specify the Macie service principal (`macie.amazonaws.com`).

To disable trusted service access by using the AWS Command Line Interface (AWS CLI), run the disable-aws-service-access command of the AWS Organizations API. For the `service-principal` parameter, specify the Macie service principal (`macie.amazonaws.com`). For example:

```
C:\> aws organizations disable-aws-service-access --service-principal
 macie.amazonaws.com
```

Managing multiple Macie accounts by invitation

> **Note**
>
> We recommend using AWS Organizations instead of Macie invitations to manage member accounts. For more information, see Managing multiple Macie accounts with AWS Organizations.

You can centrally manage multiple Amazon Macie accounts in two ways, by integrating Macie with AWS Organizations or by using membership invitations. If you use membership invitations, a designated Macie administrator can manage Macie for as many as 1,000 accounts. The administrator can also access Amazon Simple Storage Service (Amazon S3) inventory data and discover sensitive data in S3 buckets that the accounts own. For details about tasks that the administrator can perform, see Macie administrator and member account relationships.

In an invitation-based organization, you associate Macie accounts with each other by sending and accepting membership invitations in Macie. If you send an invitation and it's accepted by another account, you become the Macie administrator for the other account and the other account becomes a member account in your organization. If you receive and accept an invitation, your account becomes a member account and the Macie administrator can access certain Macie settings, data, and resources for your account.

If you create an invitation-based organization in Macie, you can subsequently transition to using AWS Organizations instead. You can also use both methods at the same time to manage multiple Macie accounts. For example, if your AWS environment includes test accounts, you might exclude the accounts from your organization in AWS Organizations and manage them separately by invitation.

The topics in this section explain how to create and participate in an invitation-based organization, and how to perform various administrative tasks for the organization.

Topics

- Creating and managing an invitation-based organization in Macie
- Reviewing Macie accounts for an invitation-based organization
- Changing the Macie administrator account for an invitation-based organization
- Managing your membership in an organization in Macie

Considerations for invitation-based organizations in Macie

> **Note**
>
> We recommend using AWS Organizations instead of Macie invitations to manage member accounts. For more information, see Managing multiple Macie accounts with AWS Organizations.

Before you create or begin managing an invitation-based organization in Amazon Macie, consider the following requirements and recommendations. Also ensure that you understand the relationship between Macie administrator and member accounts.

Topics

- Choosing a Macie administrator account
- Sending invitations and managing Macie member accounts
- Responding to and managing membership invitations
- Transitioning to AWS Organizations

Choosing a Macie administrator account

While you determine which account should be the Macie administrator account for the organization, keep the following in mind:

- An organization can have only one Macie administrator account.
- An account can't be a Macie administrator and member account at the same time.
- Macie is a Regional service. This means that the association between a Macie administrator account and a member account is Regional—the association exists only in the AWS Region that an invitation is sent from and accepted in. For example, if the Macie administrator sends

invitations in the US East (N. Virginia) Region and those invitations are accepted, the Macie administrator can manage the member accounts only in that Region.

- To centrally manage Macie accounts in multiple AWS Regions, the Macie administrator must sign in to each Region where the organization currently uses or plans to use Macie, and send invitations to the appropriate accounts in each of those Regions. For a list of Regions where Macie is currently available, see Amazon Macie endpoints and quotas in the *AWS General Reference*.
- A member account can be associated with only one Macie administrator account at a time. If your organization uses Macie in multiple Regions, this means that the Macie administrator account must be the same in all of those Regions. However, administrator and member accounts must send and accept invitations separately in each Region.

If the Macie administrator's AWS account is suspended, isolated, or closed, all associated member accounts are automatically removed as member accounts but Macie continues to be enabled for the accounts. The accounts become standalone Macie accounts. If automated sensitive data discovery was enabled for a member account, it's disabled for the account. This also disables access to statistical data, inventory data, and other information that Macie produced and directly provided while performing automated discovery for the account. After 30 days, this data expires and Macie permanently deletes it. To restore access to the data before it expires, restore the Macie administrator's AWS account, and then use that account to create and configure the organization again.

Sending invitations and managing Macie member accounts

As the Macie administrator for an invitation-based organization, keep the following in mind when you send invitations and manage accounts in the organization:

- If you send an invitation, related data might be transferred across AWS Regions. This is the case because Macie verifies the receiving account's email address by using an email verification service that operates only in the US East (N. Virginia) Region.
- You can send an invitation to any active AWS account, including accounts that haven't enabled Macie. However, to accept or decline an invitation, the receiving account must enable Macie in the Region that the invitation was sent from.
- In each AWS Region, a Macie administrator account can be associated with no more than 1,000 accounts by invitation. This includes accounts that haven't responded to invitations yet. If your account meets this quota, you can't add or invite additional accounts. To determine how

many accounts are currently associated with your account, you can use the **Accounts** page on the Amazon Macie console or the ListMembers operation of the Amazon Macie API. For more information, see Reviewing Macie accounts for an invitation-based organization.

To reduce the number of associated accounts, you can: delete associations with accounts that aren't currently member accounts, remove the necessary number of member accounts, or a combination of the two. If an account resigns from your organization or declines an invitation that you sent, it also reduces the number of accounts that are associated with your account.

- An account can be associated with only one Macie administrator account at a time. This means that an account can't accept your invitation if it's already associated with another Macie administrator account. The account must first disassociate from its current Macie administrator account.
- In an invitation-based organization, a member account can disassociate from its Macie administrator account at any time. If this happens, Macie continues to be enabled for the account but the account becomes a standalone Macie account. Macie doesn't notify you if a member account disassociates from your administrator account. However, the account continues to appear in your account inventory and it has a status of **Member resigned**.
- If you remove a member account from your organization, Macie continues to be enabled for the account. The account becomes a standalone Macie account.

Responding to and managing membership invitations

As a recipient of an invitation or a member of an invitation-based organization, keep the following in mind when you respond to and manage invitations that you receive:

- Before you accept an invitation, ensure that you understand the relationship between Macie administrator and member accounts.
- Your account can be associated with only one Macie administrator account at a time. If you accept an invitation and subsequently want to join another organization (by invitation or through AWS Organizations), you have to first disassociate your account from its current Macie administrator account. You can then join the other organization.
- To accept or decline an invitation, you have to enable Macie in the AWS Region that the invitation was sent from. The account that sent the invitation can't enable Macie in that Region for you. Declining an invitation is optional. If you decline an invitation, you can optionally disable Macie in the applicable Region after you decline the invitation.

- If you're a Macie administrator, you can't accept an invitation to become a member account—an account can't be a Macie administrator and member account at the same time. To become a member account, you must first disassociate your account from all of its member accounts by removing all member accounts from your current organization.
- Macie is a Regional service. If you accept an invitation, the association between your account and the Macie administrator account is Regional—the association exists only in the AWS Region that the invitation was sent from and accepted in.
- If you use Macie in multiple Regions, the Macie administrator account for your account has to be the same in all of those Regions. However, the Macie administrator has to send invitations to you separately in each Region, and you have to accept the invitations separately in each Region.
- You can disassociate your account from a Macie administrator account at any time. Similarly, your Macie administrator can remove your account from their organization at any time. If either happens:
  - Macie continues to be enabled for your account. Your account becomes a standalone Macie account.
  - Automated sensitive data discovery is disabled for your account, if it was enabled. This also disables access to existing statistical data, inventory data, and other information that Macie produced and directly provided while performing automated discovery for your account. You can enable automated discovery for your account again. However, this doesn't restore access to the existing data. Instead, Macie generates and maintains new data while it performs automated discovery for your account.

Transitioning to AWS Organizations

After you create an invitation-based organization in Macie, you can transition to using AWS Organizations instead. To simplify the transition, we recommend that you designate the existing, invitation-based administrator account as the Macie administrator account for the organization in AWS Organizations.

If you do this, all currently associated member accounts continue to be members. If a member account is part of the organization in AWS Organizations, the account's association automatically changes from **By invitation** to **Via AWS Organizations** in Macie. If a member account isn't part of the organization in AWS Organizations, the account's association continues to be **By invitation**. In both cases, the accounts continue to be associated with the Macie administrator account as member accounts. For sensitive data discovery, this also means that the accounts can continue to access statistical and other data that Macie produced and directly provided while performing

automated sensitive data discovery for the accounts. In addition, if the Macie administrator configured sensitive data discovery jobs to analyze data for the accounts, subsequent job runs will continue to include resources that the accounts own.

We recommend this approach because a member account can be associated with only one Macie administrator account at a time. If you designate a different account as the Macie administrator account for an organization in AWS Organizations, the designated administrator won't be able to manage accounts that are already associated with another Macie administrator account by invitation. Each member account must first disassociate from its current, invitation-based administrator account. Only then can the Macie administrator for the AWS Organizations organization add the member account to their organization and begin managing Macie for the account.

After you integrate Macie with AWS Organizations and configure your organization in Macie, you can optionally designate a different Macie administrator account for the organization. You can also continue to use invitations to associate and manage member accounts that aren't part of your organization in AWS Organizations.

For information about integrating Macie with AWS Organizations, see Managing multiple Macie accounts with AWS Organizations.

Creating and managing an invitation-based organization in Macie

> **Note**
>
> We recommend using AWS Organizations instead of Macie invitations to manage member accounts. For more information, see Managing multiple Macie accounts with AWS Organizations.

To create an invitation-based organization in Amazon Macie, you start by determining which account you want to be the Macie administrator account for the organization. You then use that account to add member accounts—you send membership invitations to other AWS accounts, inviting the accounts to join the organization as Macie member accounts in the current AWS Region. To create the organization in multiple Regions, send membership invitations from each Region in which the other accounts currently use or plan to use Macie.

When an account accepts an invitation, it becomes a Macie member account that's associated with the Macie administrator account in the applicable Region. The Macie administrator account can then access certain Macie settings, data, and resources for the member account in that Region.

As the Macie administrator for an invitation-based organization, you can review Amazon Simple Storage Service (Amazon S3) inventory data and policy findings for member accounts. You can also enable automated sensitive data discovery and run sensitive data discovery jobs to detect sensitive data in S3 buckets that member accounts own. For a detailed list of the tasks that you can perform, see Macie administrator and member account relationships.

By default, Macie gives you visibility into relevant data and resources for your organization overall. You can also drill down to review data and resources for individual accounts in your organization. For example, if you use the Summary dashboard to assess your organization's Amazon S3 security posture, you can filter the data by account. Similarly, if you monitor estimated usage costs, you can access breakdowns of estimated costs for individual member accounts.

In addition to tasks that are common to administrator and member accounts, you can centrally perform various administrative tasks for your organization. Before you perform these tasks, it's a good idea to review the considerations and recommendations for managing invitation-based organizations in Macie.

Tasks

- Adding Macie member accounts to an invitation-based organization
- Suspending Macie for member accounts in an invitation-based organization
- Removing Macie member accounts from an invitation-based organization
- Deleting associations with other accounts

Adding Macie member accounts to an invitation-based organization

As the Amazon Macie administrator for an invitation-based organization, you add member accounts to your organization by performing two primary steps:

1. Add the accounts to your account inventory in Macie. This associates the accounts with your account.
2. Send membership invitations to the accounts.

When an account accepts your invitation, it becomes a member account in your organization.

Step 1: Add the accounts

To add one or more accounts to your account inventory, you can use the Amazon Macie console or the Amazon Macie API.

Console

With the Amazon Macie console, you can add one account at a time, or add multiple accounts at the same time by uploading a comma-separated values (CSV) file. Follow these steps to add one or more accounts by using the console.

To add one account

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to add an account.
3. In the navigation pane, choose **Accounts**. The **Accounts** page opens and displays a table of the accounts that are currently associated with your account.
4. Choose **Add accounts**.
5. In the **Enter account details** section, choose **Add account**. Then do the following:
   - For **Account ID**, enter the 12-digit account ID for the AWS account to add.
   - For **Email address**, enter the email address for the AWS account to add.
6. Choose **Add**.
7. At the bottom of the page, choose **Next**.

Macie adds the account to your account inventory. The account's type is **By invitation** and its status is **Created**. To add the account in additional Regions, repeat the preceding steps in each additional Region.

To add multiple accounts

1. By using a text editor, create a CSV file as follows:
   a. Add the following header as the first line of the file: `Account ID,Email`
   b. For each account, create a new line that has the 12-digit account ID for the AWS account to add and the email address for the account. Separate the entries with a comma, for example: `111111111111,janedoe@example.com`

The email address must match the email address that's associated with the AWS account.

c. Verify that the file's contents are formatted as shown in the following example, which contains the required header and information for three accounts:

```
Account ID,Email
111111111111,janedoe@example.com
222222222222,jorgesouza@example.com
333333333333,lijuan@example.com
```

d. Save the file on your computer.

2. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
3. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to add the accounts.
4. In the navigation pane, choose **Accounts**. The **Accounts** page opens and displays a table of the accounts that are currently associated with your account.
5. Choose **Add accounts**.
6. In the **Enter account details** section, choose **Upload list (CSV)**.
7. Choose **Browse**, and then select the CSV file that you created in step 1.
8. Choose **Add accounts**.
9. At the bottom of the page, choose **Next**.

Macie adds the accounts to your account inventory. Their type is **By invitation** and their status is **Created**. To add the accounts in additional Regions, repeat steps 3 through 8 in each additional Region.

API

To add one or more accounts programmatically, use the CreateMember operation of the Amazon Macie API. When you submit your request, use the supported parameters to specify the 12-digit account ID and email address for each AWS account to add. Also specify the Region that the request applies to. To add accounts in additional Regions, submit the request in each additional Region.

To add accounts by using the AWS Command Line Interface (AWS CLI), run the create-member command. Use the `region` parameter to specify the Region in which to add the accounts. Use

the `account` parameters to specify the account ID and email address for each AWS account to add. For example:

```
C:\> aws macie2 create-member --region us-east-1 --account={\"accountId\":
\"111111111111\",\"email\":\"janedoe@example.com\"}
```

Where *us-east-1* is the Region in which to add the account (the US East (N. Virginia) Region) and the `account` parameters specify the account ID (*111111111111*) and email address (*janedoe@example.com*) for the account to add.

If your request succeeds, Macie adds each account to your account inventory with a status of `Created` and you receive output similar to the following:

```
{
    "arn": "arn:aws:macie2:us-east-1:123456789012:member/111111111111"
}
```

Where `arn` is the Amazon Resource Name (ARN) of the resource that was created for the association between your account and the account that you added. In this example, `123456789012` is the account ID for the account that created the association and `111111111111` is the account ID for the account that was added.

Step 2: Send membership invitations to the accounts

After you add an account to your account inventory, you can invite the account to join your organization as a Macie member account. To do this, send a membership invitation to the account. When you send an invitation, an **Accounts** badge and notification appear on the Amazon Macie console for the recipient's account, if Macie is enabled for the account. Macie also creates an AWS Health event for the account.

Depending on whether you use the Amazon Macie console or API to send the invitation, Macie also sends the invitation to the email address that you specified for the recipient's account when you added the account. The email message indicates that you would like to become the Macie administrator for their account, and it includes the account ID for your AWS account and the recipient's AWS account. The message also explains how to access the invitation. You can optionally add custom text to the message.

To send a membership invitation to one or more accounts, you can use the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to send a membership invitation by using the Amazon Macie console.

To send a membership invitation

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to send the invitation.
3. In the navigation pane, choose **Accounts**. The **Accounts** page opens and displays a table of the accounts that are currently associated with your account.
4. In the **Existing accounts** table, select the checkbox for each account that you want to send the invitation to.

   > **Tip**
   >
   > To more easily identify accounts that you added and haven't sent invitations to yet, you can filter the table. To do this, place your cursor in the filter box above the table, and then choose **Status**. Then choose **Status = Created**.

5. On the **Actions** menu, choose **Invite**.
6. (Optional) In the **Message** box, enter any custom text that you want to include in the email message that contains the invitation. The text can contain as many as 80 alphanumeric characters.
7. Choose **Invite**.

To send the invitation in additional AWS Regions, repeat the preceding steps in each additional Region.

After you send the invitation, the status of a recipient account changes to **Email verification in progress** in your account inventory. If Macie can verify an account's email address, the account's status subsequently changes to **Invited**. If Macie can't verify the address, the account's status changes to **Email verification failed**. If this happens, work with the account owner to get the correct email address. Then delete the association between your accounts, add the account again, and send the invitation again.

When a recipient accepts an invitation, the status of the recipient's account changes to **Enabled** in your account inventory. If a recipient declines an invitation, the recipient's account is disassociated from your account and removed from your account inventory.

API

To send an invitation programmatically, use the CreateInvitations operation of the Amazon Macie API. When you submit your request, use the supported parameters to specify the 12-digit account ID for each AWS account to send the invitation to. An account ID must match the account ID for an account in your account inventory. Otherwise, an error occurs. Also specify the Region to send the invitation from. To send the invitation from additional Regions, submit the request in each additional Region.

In your request, you can also specify whether to send the invitation as an email message, and whether to include custom text in that message. If you choose to send an email message, Macie sends the invitation to the email address that you specified for an account when you added the account to your account inventory. To send the invitation as an email message, omit the `disableEmailNotification` parameter or set the value for the parameter to `false`. (The default value is `false`.) To add custom text to the message, use the `message` parameter to specify the text to add. The text can contain as many as 80 alphanumeric characters.

To send invitations by using the AWS CLI, run the create-invitations command. Use the `region` parameter to specify the Region to send the invitation from. Use the `account-ids` parameter to specify the account ID for each AWS account to send the invitation to. For example:

```
C:\> aws macie2 create-invitations --region us-east-1 --account-
ids=[\"111111111111\",\"222222222222\",\"333333333333\"]
```

Where *us-east-1* is the Region to send the invitation from (the US East (N. Virginia) Region) and the `account-ids` parameter specifies account IDs for three accounts to send the invitation to. To send an invitation as an email message too, also include the `no-disable-email-notification` parameter and optionally include the `message` parameter to specify custom text to add to the message.

After you send the invitation, the status of each recipient account changes to `EmailVerificationInProgress`. If Macie can verify an account's email address, the account's status subsequently changes to `Invited`. If Macie can't verify the address, the account's status changes to `EmailVerificationFailed`. If this happens, work with the

account owner to get the correct address. Then delete the association between your accounts, add the account again, and send the invitation again.

When a recipient accepts an invitation, the status of the recipient's account changes to `Enabled` in your account inventory. If a recipient declines an invitation, the recipient's account is disassociated from your account and removed from your account inventory.

Suspending Macie for member accounts in an invitation-based organization

As the Amazon Macie administrator for an organization, you can suspend Macie in a specific AWS Region for individual member accounts in your organization. Note, however, that you can't re-enable Macie for a member account after you suspend it. Only a user of the account can subsequently re-enable Macie for the account.

When you suspend Macie for a member account:

- Macie loses access to and stops providing metadata about the account's Amazon S3 data in the Region.
- Macie stops performing all activities for the account in the Region. This includes monitoring S3 buckets for security and access control, performing automated sensitive data discovery, and running sensitive data discovery jobs that are currently in progress.
- Macie cancels all sensitive data discovery jobs that were created by the account in the Region. A job can't be resumed or restarted after it's cancelled. If you created jobs to analyze data that the member account owns, Macie doesn't cancel your jobs. Instead, the jobs skip resources that are owned by the account.

While it's suspended, Macie retains the Macie session identifier, settings, and resources that it stores or maintains for the account in the applicable Region. Macie also retains certain data for the account in the Region. For example, the account's findings remain intact and aren't affected for up to 90 days. If automated sensitive data discovery was enabled for the account, existing results also remain intact and aren't affected for up to 30 days. The account isn't charged for using Macie in the applicable Region while Macie is suspended for the account in that Region.

To suspend Macie for a member account in an invitation-based organization

To suspend Macie for a member account in an invitation-based organization, you can use the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to suspend Macie for a member account by using the Amazon Macie console.

To suspend Macie for a member account

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to suspend Macie for a member account.
3. In the navigation pane, choose **Accounts**. The **Accounts** page opens and displays a table of the accounts that are currently associated with your account.
4. In the **Existing accounts** table, select the checkbox for the account that you want to suspend Macie for.
5. On the **Actions** menu, choose **Suspend Macie**.
6. Confirm that you want to suspend Macie for the selected account.

After you confirm the suspension, the status of the account changes to **Paused (suspended)** in your account inventory.

To suspend Macie for the account in additional Regions, repeat the preceding steps in each additional Region.

API

To suspend Macie for a member account programmatically, use the UpdateMemberSession operation of the Amazon Macie API. When you submit your request, use the `id` parameter to specify the 12-digit account ID of the AWS account that you want to suspend Macie for. For the `status` parameter, specify `PAUSED` as the new status for Macie. Also specify the Region that the request applies to. To suspend Macie in additional Regions, submit your request in each additional Region.

To retrieve the account ID for the member account, you can use the ListMembers operation of the Amazon Macie API. If you do this, consider filtering the results by including the `onlyAssociated` parameter in your request. If you set this parameter's value to `true`, Macie returns a `members` array that provides details about only those accounts that are currently member accounts for your administrator account.

To suspend Macie for a member account by using the AWS CLI, run the update-member-session command. Use the `region` parameter to specify the Region in which to suspend Macie. Use the

`id` parameter to specify the account ID for the account to suspend Macie for. For the `status` parameter, specify PAUSED. For example:

```
C:\> aws macie2 update-member-session --region us-east-1 --id 123456789012 --status
 PAUSED
```

Where *us-east-1* is the Region in which to suspend Macie (the US East (N. Virginia) Region), *123456789012* is the account ID for the account to suspend Macie for, and PAUSED is the new status of Macie for the account.

If your request succeeds, Macie returns an empty response and the status of the specified account changes to `Paused` in your account inventory.

Removing Macie member accounts from an invitation-based organization

As an Amazon Macie administrator, you can remove a member account from your organization. You do this by disassociating the account from your Macie administrator account.

If you remove a member account, Macie continues to be enabled for the account and the account continues to appear in your account inventory. However, the account becomes a standalone Macie account. Macie doesn't notify the account's owner when you remove the account. Therefore, consider contacting the account owner to ensure that they begin managing settings and resources for their account.

When you remove a member account, you lose access to all Macie settings, resources, and data for the account. This includes policy findings and metadata for S3 buckets that the account owns. In addition, you can no longer use Macie to discover sensitive data in S3 buckets that the account owns. If you already created sensitive data discovery jobs to do this, the jobs skip buckets that the account owns. If you enabled automated sensitive data discovery for the account, both you and the account lose access to statistical data, inventory data, and other information that Macie produced and directly provided while performing automated discovery for the account.

After you remove a member account, you can subsequently add it to your organization again by sending a new invitation to the account. If the account accepts the new invitation and you enable automated sensitive data discovery for it within 30 days, you also regain access to data and information that Macie previously produced and directly provided while performing automated discovery for the account. In addition, subsequent runs of your existing jobs start including the account's S3 buckets again.

If you remove a member account and don't plan to add it again, you can remove it from your account inventory completely. To learn how, see Deleting associations with other accounts.

To remove a member account from an invitation-based organization

To remove a member account from your organization, you can use the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to remove a member account by using the Amazon Macie console.

To remove a member account

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to remove a member account.
3. In the navigation pane, choose **Accounts**. The **Accounts** page opens and displays a table of the accounts that are currently associated with your account.
4. In the **Existing accounts** table, select the checkbox for the account that you want to remove.
5. On the **Actions** menu, choose **Disassociate account**.
6. Confirm that you want to remove the selected account as a member account.

After you confirm your selection, the status of the account changes to **Removed (disassociated)** in your account inventory.

To remove the member account in additional Regions, repeat the preceding steps in each additional Region.

API

To remove a member account programmatically, use the DisassociateMember operation of the Amazon Macie API. When you submit your request, use the `id` parameter to specify the 12-digit AWS account ID for the member account to remove. Also specify the Region that the request applies to. To remove the account in additional Regions, submit your request in each additional Region.

To retrieve the account ID for the account to remove, you can use the ListMembers operation of the Amazon Macie API. If you do this, consider filtering the results by including the

`onlyAssociated` parameter in your request. If you set this parameter's value to `true`, Macie returns a `members` array that provides details about only those accounts that are currently member accounts for your account.

To remove a member account by using the AWS CLI, run the disassociate-member command. Use the `region` parameter to specify the Region in which to remove the account. Use the `id` parameter to specify the account ID for the account to remove. For example:

```
C:\> aws macie2 disassociate-member --region us-east-1 --id 123456789012
```

Where *us-east-1* is the Region in which to remove the account (the US East (N. Virginia) Region) and *123456789012* is the account ID for the account to remove.

If your request succeeds, Macie returns an empty response and the status of the specified account changes to `Removed` in your account inventory.

Deleting associations with other accounts

After you add an account to your account inventory in Amazon Macie, you can delete the association between your account and the other account. You can do this for any account in your inventory except:

- An account that's part of your organization in AWS Organizations. This type of association is controlled through AWS Organizations not Macie.
- A member account that accepted a Macie membership invitation to join your organization. If this is the case, you must remove the member account before you can delete the association.

When you delete an association, Macie removes the account from your account inventory. If you want to subsequently restore the association, you have to add the account again as if it were a completely new account.

To delete an association with another account

To delete an association between your account and another account, you can use the Amazon Macie console or the Amazon Macie API.

Console

To use the Amazon Macie console to delete an association with another account, follow these steps.

To delete an association

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to delete an association.
3. In the navigation pane, choose **Accounts**. The **Accounts** page opens and displays a table of the accounts that are currently associated with your account.
4. In the **Existing accounts** table, select the checkbox for the account whose association you want to delete.
5. On the **Actions** menu, choose **Delete**.
6. Confirm that you want to delete the selected association.

To delete the association in additional Regions, repeat the preceding steps in each additional Region.

API

To delete an association with another account programmatically, use the DeleteMember operation of the Amazon Macie API. When you submit your request, use the `id` parameter to specify the 12-digit account ID for the AWS account to delete the association with. Also specify the Region that the request applies to. To delete the association in additional Regions, submit your request in each additional Region.

To retrieve the account ID for the account, you can use the ListMembers operation of the Amazon Macie API. If you do this, include the `onlyAssociated` parameter in your request and set the parameter's value to `false`. If the operation is successful, Macie returns a `members` array that provides details about all the accounts that are associated with your account, including accounts that aren't currently member accounts.

To delete an association with another account by using the AWS CLI, run the delete-member command. Use the `region` parameter to specify the Region in which to delete the association. Use the `id` parameter to specify the account ID for the account. For example:

```
C:\> aws macie2 delete-member --region us-east-1 --id 123456789012
```

Where *us-east-1* is the Region in which to delete the association with the other account (the US East (N. Virginia) Region) and *123456789012* is the account ID for the account.

If your request succeeds, Macie returns an empty response and the association between your account and the other account is deleted. The previously associated account is removed from your account inventory.

Reviewing Macie accounts for an invitation-based organization

> **Note**
>
> We recommend using AWS Organizations instead of Macie invitations to manage member accounts. For more information, see Managing multiple Macie accounts with AWS Organizations.

If you're the Amazon Macie administrator for an invitation-based organization, Macie provides you with an inventory of the accounts that are associated with your Macie account in each AWS Region where you use Macie. You can use this inventory to review account statistics and details for your organization. You can also use it to perform certain management tasks for member accounts, and manage the status of the relationship between your account and other accounts.

To review accounts for an invitation-based organization

To review the accounts in your organization, you can use the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to review your organization's accounts by using the Amazon Macie console.

To review your organization's accounts

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to review your organization's accounts.
3. In the navigation pane, choose **Accounts**.

The **Accounts** page opens and displays aggregated statistics and a table of the accounts that are associated with your Macie account in the current AWS Region.

At the top of the **Accounts** page, you'll find the following aggregated statistics.

Via AWS Organizations

If you're the Macie administrator for an organization in AWS Organizations, **Active** reports the total number of accounts that are associated with your account through AWS Organizations and are currently Macie member accounts in your organization. Macie is enabled for these accounts and you're the Macie administrator of the accounts.

All reports the total number of accounts that are associated with your account through AWS Organizations. This includes accounts that aren't currently Macie member accounts. It also includes member accounts that Macie is currently suspended for.

By invitation

Active reports the total number of accounts that are currently Macie member accounts in your invitation-based organization. Macie is enabled for these accounts and you're the Macie administrator of the accounts because they accepted a membership invitation from you.

All reports the total number of accounts that are associated with your account by Macie invitation, including accounts that haven't responded to an invitation from you.

Active/All

Active reports the total number of accounts that Macie is currently enabled for in your organization, including your own account. You're the Macie administrator of these accounts through AWS Organizations or by Macie invitation.

All reports the total number of accounts that are associated with your account, through AWS Organizations or by invitation, plus your own account. This includes accounts that haven't responded to a Macie membership invitation from you. It also includes accounts that are associated with your account through AWS Organizations and aren't currently Macie member accounts.

In the table, you'll find details about each account in the current Region. The table includes all the accounts that are associated with your Macie account by Macie invitation or through AWS Organizations.

Account ID

The account ID and email address for the AWS account.

Name

The account name for the AWS account. This value is typically **N/A** for your own account, and accounts that are associated with your account by invitation.

Type

How the account is associated with your account, by invitation or through AWS Organizations. For your own account, this value is **Current account**.

Status

The status of the relationship between your account and the account. For an account in an invitation-based organization (**Type** is **By invitation**), possible values are:

- **Account suspended** – The AWS account is suspended.
- **Created (Invite)** – You added the account but haven't sent a membership invitation to it.
- **Email verification failed** – You tried to send a membership invitation to the account but the specified email address isn't valid for the account.
- **Email verification in progress** – You sent a membership invitation to the account and Macie is processing the request.
- **Enabled** – The account is a member account. Macie is enabled for the account and you're the Macie administrator of the account.
- **Invited** – You sent a membership invitation to the account and the account hasn't responded to your invitation.
- **Member resigned** – The account was previously a member account. However, the account resigned from your organization by disassociating from your account.
- **Paused (suspended)** – The account is a member account but Macie is currently suspended for the account.
- **Region disabled** – The current Region is disabled for the AWS account.
- **Removed (disassociated)** – The account was previously a member account. However, you removed it as a member account by disassociating it from your account.

Last status update

When you or the associated account most recently performed an action that affected the relationship between your accounts.

Automated sensitive data discovery

Whether automated sensitive data discovery is currently enabled or disabled for the account.

To sort the table by a specific field, choose the column heading for the field. To change the sort order, choose the column heading again. To filter the table, place your cursor in the filter box, and then add a filter condition for a field. To further refine the results, add filter conditions for additional fields.

API

To review your organization's accounts programmatically, use the ListMembers operation of the Amazon Macie API and specify the Region that your request applies to. To review the details in additional Regions, submit your request in each additional Region.

When you submit your request, use the `onlyAssociated` parameter to specify which accounts to include in the response. By default, Macie returns details about only those accounts that are member accounts in the specified Region, by invitation or through AWS Organizations. To retrieve the details of all associated accounts, including accounts that aren't member accounts, include the `onlyAssociated` parameter in your request and set the parameter's value to `false`.

To review your organization's accounts by using the AWS Command Line Interface (AWS CLI), run the list-members command. For the `only-associated` parameter, specify whether to include all associated accounts or only member accounts. To include only member accounts, omit this parameter or set the parameter's value to `true`. To include all accounts, set this value to `false`. For example:

```
C:\> aws macie2 list-members --region us-east-1 --only-associated false
```

Where *us-east-1* is the Region that the request applies to, the US East (N. Virginia) Region.

If your request succeeds, Macie returns a `members` array. The array contains a `member` object for each account that meets the criteria specified in the request. In that object, the `relationshipStatus` field indicates the current status of the association between your account and the other account in the specified Region. For an account in an invitation-based organization, possible values are:

- `AccountSuspended` – The AWS account is suspended.

- `Created` – You added the account but haven't sent a membership invitation to it.
- `EmailVerificationFailed` – You tried to send a membership invitation to the account but the specified email address isn't valid for the account.
- `EmailVerificationInProgress` – You sent a membership invitation to the account and Macie is processing the request.
- `Enabled` – The account is a member account. Macie is enabled for the account and you're the Macie administrator of the account.
- `Invited` – You sent a membership invitation to the account and the account hasn't responded to your invitation.
- `Paused` – The account is a member account but Macie is currently suspended (paused) for the account.
- `RegionDisabled` – The current Region is disabled for the AWS account.
- `Removed` – The account was previously a member account. However, you removed it as a member account by disassociating it from your account.
- `Resigned` – The account was previously a member account. However, the account resigned from your organization by disassociating from your account.

For information about other fields in the `member` object, see Members in the *Amazon Macie API Reference*.

Changing the Macie administrator account for an invitation-based organization

> **Note**
>
> We recommend using AWS Organizations instead of Macie invitations to manage member accounts. For more information, see Managing multiple Macie accounts with AWS Organizations.

After you create and establish an invitation-based organization, you can change the Amazon Macie administrator account for the organization. To do this, administrators and members of the organization should take the following steps:

1. The current Macie administrator optionally exports the current inventory of member accounts for the organization. This simplifies the transition by helping you identify accounts that should continue to be part of the organization.
2. The current Macie administrator removes all member accounts from the current organization. This disassociates the accounts from the current administrator account. Macie continues to be enabled for the accounts but the accounts become standalone Macie accounts.

> **Important**
>
> When the current Macie administrator removes the member accounts, Macie automatically disables automated sensitive data discovery for the accounts. This also disables access to statistical data, inventory data, and other information that Macie produced and directly provided while performing automated discovery for the accounts. When the transition to the new organization is complete, the new Macie administrator can't access this data.

3. The new Macie administrator adds the previous member accounts to the new organization. This associates the accounts with the new administrator account.
4. Each member account accepts the invitation to join the new organization. When an account accepts the invitation, the account becomes a member account in the new organization. The new Macie administrator can then access Macie settings, data, and resources for the account. If automated sensitive data discovery was previously enabled for the account, this doesn't include data that Macie previously produced and directly provided while performing automated discovery for the account. Instead, Macie generates and maintains new data for the account, if the new Macie administrator enables automated discovery for the account.

If your organization uses Macie in multiple AWS Regions, perform the preceding steps in each of those Regions.

To export the current inventory of member accounts, the current Macie administrator can use the Amazon Macie console or the Amazon Macie API. With the console, the current administrator can export the data to a comma-separated values (CSV) file. The new administrator can then use the console to upload the CSV file and add all the accounts (in bulk) to the new organization.

To export member account data by using the console

1. Sign in to the AWS Management Console using the current Macie administrator account.

2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to export the data.
3. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
4. In the navigation pane, choose **Accounts**. The **Accounts** page opens and displays a table of the accounts that are associated with the current Macie administrator account.
5. (Optional) To filter the table and show only those accounts that are currently member accounts in the organization, use the filter box above the table to add the following filter conditions:

   - **Type** = **Invitation**
   - **Status** = **Enabled**
   - **Status** = **Paused**
6. In the table, select the checkbox for each member account to include in the exported data.
7. Choose **Export CSV**.
8. Specify a name and location for the file.

With the Amazon Macie API, the current Macie administrator can retrieve the data in JSON format. The new Macie administrator can then use that data to generate the list of account IDs and email addresses for the accounts to add and invite to the new organization. To retrieve the data in JSON format, use the ListMembers operation of the Amazon Macie API. If the operation succeeds, Macie returns a `members` array that provides details about all the accounts that are associated with the administrator's account. If an account is currently a member account, the value for the `relationshipStatus` property of the account is `Enabled` or `Paused`, and the `invitedAt` property specifies a date and time.

Managing your membership in an organization in Macie

> **Note**
>
> We recommend using AWS Organizations instead of Macie invitations to centrally manage Macie for multiple accounts. For more information, see Managing multiple Macie accounts with AWS Organizations.

If you're invited to join an organization in Amazon Macie, you can optionally accept or decline the invitation. In Macie, an organization is a set of accounts that are centrally managed as a group of

related accounts. An organization consists of one designated Macie administrator account and one or more associated member accounts.

If you accept an invitation, your account becomes a member account in the organization. When you accept, the account that sent the invitation becomes the Macie administrator account for your account—you associate your account with the other account and you enable an administrator-member relationship between the accounts. The Macie administrator account can then access certain Macie settings, data, and resources for your account in the applicable AWS Region. For details about tasks that the administrator account can perform, see Macie administrator and member account relationships.

If you decline an invitation, the current status and settings for your Macie account aren't changed.

Topics

Responding to membership invitations for organizations

When you receive an invitation to join an organization, Amazon Macie notifies you in several ways. By default, Macie sends the invitation to you as an email message. Macie also creates an AWS Health event for your AWS account. If you already use Macie in the AWS Region from which the invitation was sent, Macie also displays an **Accounts** badge and notification on the Macie console.

After you receive an invitation, you can optionally accept or decline the invitation. Before you respond, note the following:

- You can be a member of only one organization at a time. If you receive multiple invitations, you can accept only one. Or, if you're already a member of an organization, you have to disassociate your account from its current Macie administrator account before you can join a different organization.
- If you use Macie in multiple Regions, your account has to have the same Macie administrator account in all of those Regions. The Macie administrator has to send invitations to you separately from each Region, and you have to accept the invitations separately in each Region.
- To accept or decline an invitation, you have to enable Macie in the Region that the invitation was sent from. Declining an invitation is optional. If you enable Macie to decline an invitation, you can disable Macie in the Region after you decline the invitation. This helps ensure that you don't incur unnecessary charges for using Macie in the Region.

- If automated sensitive data discovery is enabled for your account and you accept an invitation, you lose access to statistical data, inventory data, and other information that Macie produced and directly provided while performing automated discovery for your account. After you accept an invitation, your Macie administrator can enable automated discovery for your account. However, this doesn't restore access to the existing data. Instead, Macie generates and maintains new data while it performs automated discovery for your account.

For additional considerations, see Responding to and managing membership invitations.

To respond to a membership invitation for an organization

To respond to a membership invitation, you can use the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to respond to a membership invitation by using the Amazon Macie console.

To respond to a membership invitation

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you received the invitation.
3. If you haven't enabled Macie in the Region, choose **Get started**, and then choose **Enable Macie**. You have to enable Macie before you can accept or decline an invitation.
4. In the navigation pane, choose **Accounts**.
5. Under **Administrator account**, do one of the following:
   - To accept the invitation, turn on **Accept** () next to the invitation. Then choose **Accept invitation** or **Update**, depending on whether you previously accepted another invitation.
   - To decline the invitation, choose **Decline invitation** next to the invitation, and then confirm that you want to decline the invitation.

If you received and want to respond to the invitation in additional Regions, repeat the preceding steps in each additional Region.

API

To respond to an invitation programmatically, use the AcceptInvitation or DeclineInvitations operation of the Amazon Macie API, depending on whether you want to accept or decline the invitation. When you submit your request, be sure to specify the Region that the invitation was sent from. To respond to the invitation in additional Regions, submit your request in each additional Region.

In an `AcceptInvitation` request, use the `administratorAccountId` parameter to specify the 12-digit account ID for the AWS account that sent the invitation. Use the `invitationId` parameter to specify the unique ID for the invitation to accept.

In a `DeclineInvitations` request, use the `accountIds` parameter to specify the 12-digit account ID for the AWS account that sent the invitation to decline.

To retrieve the IDs, you can use the ListInvitations operation of the Amazon Macie API. If the operation succeeds, Macie returns an `invitations` array that provides details about invitations that you've received, including the account ID for the account that sent each invitation and the unique ID for each invitation. If the value for the `relationshipStatus` property of an invitation is `Invited`, you haven't responded to the invitation yet.

To respond to an invitation by using the AWS Command Line Interface (AWS CLI), run the accept-invitation or decline-invitations command, depending on whether you want to accept or decline the invitation. Use the `region` parameter to specify the Region that the invitation was sent from. For example:

```
C:\> aws macie2 accept-invitation --region us-east-1 --administrator-account-
id 123456789012 --invitation-id d8bdad0e203fd1242e0a4721bexample
```

Where *us-east-1* is the Region that the invitation was sent from (the US East (N. Virginia) Region), *123456789012* is the account ID for the account that sent the invitation, and *d8bdad0e203fd1242e0a4721bexample* is the unique ID for the invitation to accept.

If a request to accept an invitation succeeds, Macie returns an empty response. If a request to decline an invitation succeeds, Macie returns an empty `unprocessedAccounts` array.

After you decline an invitation, the invitation persists as a resource for your Macie account. You can optionally delete it by using the DeleteInvitations operation or, for the AWS CLI, the delete-invitations command.

Disassociating from a Macie administrator account

If you accept an invitation to join an organization in Amazon Macie, you can subsequently resign from the organization by disassociating your account from its current Macie administrator account. Note that you can't do this if your account is a member account in an AWS Organizations organization. To resign from an AWS Organizations organization, work with your Macie administrator to remove your account as a Macie member account.

If you disassociate your account from its Macie administrator account, the Macie administrator loses access to all settings, data, and resources for your Macie account. This includes metadata and policy findings for Amazon S3 data that you own. This also means that the administrator can no longer analyze your Amazon S3 data by performing automated sensitive data discovery or running sensitive data discovery jobs.

When you disassociate your account, Macie continues to be enabled for your account in the applicable Region. However, your account becomes a standalone Macie account in the Region. The status of your account changes to **Member resigned** in the administrator's account inventory.

To disassociate from a Macie administrator account

To disassociate your account from its current Macie administrator account, you can use the Amazon Macie console or the Amazon Macie API.

Console

Follow these steps to disassociate your account from its Macie administrator account by using the Amazon Macie console.

To disassociate from an administrator account

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to disassociate your account from its administrator account.
3. In the navigation pane, choose **Accounts**.
4. Under **Administrator account**, turn off **Accept** () next to the invitation, and then choose **Update**.

The account continues to appear on the **Accounts** page. If you decide to re-join the organization, you can use this page to accept the original invitation again. Alternatively, you can decline and delete the invitation, which also deletes the association between your account and the other account. To do this, choose **Decline invitation**.

If you want to disassociate your account from its Macie administrator account in additional Regions, repeat the preceding steps in each additional Region.

API

To disassociate your account from its Macie administrator account programmatically, use the DisassociateFromAdministratorAccount operation of the Amazon Macie API. When you submit your request, be sure to specify the Region that the request applies to. To disassociate from the account in additional Regions, submit your request in each additional Region.

To disassociate your account from its Macie administrator account by using the AWS CLI, run the disassociate-from-administrator-account command. Use the `region` parameter to specify the Region in which to disassociate from the account.

If your request succeeds, Macie returns an empty response.

After you disassociate from the account, the original invitation persists as a resource for your Macie account unless you delete it. If you decide to re-join the organization, you can use this resource to accept the original invitation again. Alternatively, you can delete the invitation by using the DeleteInvitations operation or, for the AWS CLI, the delete-invitations command. If you delete the invitation, you also delete the association between your account and the other account.

Tagging Macie resources

A *tag* is a label that you can define and assign to AWS resources, including certain types of Amazon Macie resources. Tags can help you identify, categorize, and manage resources in different ways, such as by purpose, owner, environment, or other criteria. For example, you can use tags to: apply policies, allocate costs, distinguish between versions of resources, or identify resources that support certain compliance requirements or workflows.

You can assign tags to the following types of Macie resources: allow lists, custom data identifiers, filter rules and suppression rules for findings, and sensitive data discovery jobs. If you're the Macie administrator for an organization, you can also assign tags to member accounts in your organization.

A resource can have as many as 50 tags. Each tag consists of a required *tag key* and an optional *tag value*. A *tag key* is a general label that acts as a category for a more specific tag value. A *tag value* acts as a descriptor for a tag key.

For example, if you create custom data identifiers and sensitive data discovery jobs to analyze data at different points in a workflow (one set for staged data and another for production data), you might assign a `Stack` tag key to those resources. The tag value for this tag key might be `Staging` for custom data identifiers and jobs that analyze staged data, and `Production` for the others.

Topics

Tagging fundamentals for Macie resources

To identify, categorize, and manage Amazon Macie resources for your account, you can assign tags to the resources. A *tag* is a label that you define and assign to AWS resources, including certain types of Macie resources. Each tag consists of a required *tag key* and an optional *tag value*. A *tag key* is a general label that acts as a category for a more specific tag value. A *tag value* acts as a descriptor for a tag key. A resource can have as many as 50 tags.

You can assign tags to the following types of Macie resources:

- Allow lists
- Custom data identifiers
- Filter rules and suppression rules for findings
- Sensitive data discovery jobs

If you're the Macie administrator for an organization, you can also assign tags to member accounts in your organization.

By assigning tags to Macie resources, you can identify and manage the resources in different ways, such as by purpose, owner, environment, or other criteria. This can help you perform tasks such as apply policies, allocate costs, distinguish between resources, or identify resources that support certain compliance requirements or workflows. For example, if you create custom data identifiers and sensitive data discovery jobs to analyze data at different points in a workflow (one set for staged data and another for production data), you might assign a `Stack` tag key to those resources. The tag value for this tag key might be `Staging` for custom data identifiers and jobs that analyze staged data, and `Production` for the others.

As you define and assign tags to Macie resources, keep the following in mind:

- Each resource can have a maximum of 50 tags.
- For each resource, each tag key must be unique and it can have only one tag value.
- Tag keys and values are case sensitive. As a best practice, we recommend that you define a strategy for capitalizing tags and implement that strategy consistently across your resources.
- A tag key can have a maximum of 128 UTF-8 characters. A tag value can have a maximum of 256 UTF-8 characters. The characters can be letters, numbers, spaces, or the following symbols: _ . : / = + - @
- The `aws:` prefix is reserved for use by AWS. You can't use it in any tag keys or values that you define. In addition, you can't change or remove tag keys or values that use this prefix. Tags that use this prefix don't count against the quota of 50 tags for a resource.
- Any tags that you assign are available only for your AWS account and only in the AWS Region in which you assign them.
- If you delete a resource, any tags that are assigned to the resource are also deleted.

For additional restrictions, tips, and best practices, see the Tagging AWS Resources User Guide.

> **Important**
>
> Do not store confidential or other types of sensitive data in tags. Tags are accessible from many AWS services, including AWS Billing and Cost Management. They aren't intended to be used for sensitive data.

To add and manage tags for Macie resources, you can use Macie or AWS Resource Groups. AWS Resource Groups is a service that's designed to help you group and manage AWS resources as a single unit instead of individually. If you use Macie, you can add tags to a resource when you create the resource. You can also add and manage tags for individual existing resources. If you use AWS Resource Groups, you can add and manage tags in bulk for multiple existing resources spanning multiple AWS services, including Macie. For more information, see the Tagging AWS Resources User Guide.

Adding tags to Macie resources

A *tag* is a label that you can define and assign to AWS resources, including certain types of Amazon Macie resources. By using tags, you can identify, categorize, and manage resources in different ways, such as by purpose, owner, environment, or other criteria. For example, you can use tags to: apply policies, allocate costs, distinguish between versions of resources, or identify resources that support certain compliance requirements or workflows.

You can add tags to the following types of Macie resources:

- Allow lists
- Custom data identifiers
- Filter rules and suppression rules for findings
- Sensitive data discovery jobs

If you're the Macie administrator for an organization, you can also add tags to member accounts in your organization.

A resource can have as many as 50 tags. Each tag consists of a required *tag key* and an optional *tag value*. A *tag key* is a general label that acts as a category for a more specific tag value. A *tag value* acts as a descriptor for a tag key. For more information about tagging options and requirements, see Tagging fundamentals.

You can add tags to Macie resources in several ways. You can use Macie directly. You can also use the Tag Editor on the AWS Resource Groups console or tagging operations of the AWS Resource Groups Tagging API. AWS Resource Groups is a service that's designed to help you group and manage AWS resources as a single unit instead of individually. If you use Macie, you can add tags to a resource when you create the resource. You can also add tags to individual existing resources. With AWS Resource Groups, you can add tags in bulk for multiple existing resources spanning multiple AWS services, including Macie.

To add tags to a Macie resource

To add tags to an individual Macie resource, you can use the Amazon Macie console or the Amazon Macie API. To add tags to multiple Macie resources at the same time, use the AWS Resource Groups console or the AWS Resource Groups Tagging API. For more information, see the Tagging AWS Resources User Guide.

> **Important**
>
> Adding tags to a resource can affect access to the resource. Before you add a tag to a resource, review any AWS Identity and Access Management (IAM) policies that might use tags to control access to resources. For more information, see Controlling access to AWS resources using tags in the *IAM User Guide*.

Console

When you create an allow list, custom data identifier, or sensitive data discovery job, the Amazon Macie console provides options for adding tags to the resource. Follow the instructions on the console to add tags to these types of resources when you create the resources. To add tags to a filter rule, suppression rule, or member account, you have to create the resource before you can add tags to it.

To add one or more tags to an existing resource by using the Amazon Macie console, follow these steps.

To add a tag to a resource

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. Depending on the type of resource that you want to add a tag to, do one of the following:

- For an allow list, choose **Allow lists** in the navigation pane. In the table, select the checkbox for the list. Then choose **Manage tags** on the **Actions** menu.
- For a custom data identifier, choose **Custom data identifiers** in the navigation pane. In the table, select the checkbox for the custom data identifier. Then choose **Manage tags** on the **Actions** menu.
- For a filter or suppression rule, choose **Findings** in the navigation pane. In the **Saved rules** list, choose the edit icon () next to the rule. Then choose **Manage tags**.
- For a member account in your organization, choose **Accounts** in the navigation pane. In the table, select the checkbox for the account. Then choose **Manage tags** on the **Actions** menu.
- For a sensitive data discovery job, choose **Jobs** in the navigation pane. In the table, select the checkbox for the job. Then choose **Manage tags** on the **Actions** menu.

The **Manage tags** window lists all the tags that are currently assigned to the resource.

3. In the **Manage tags** window, choose **Edit tags**.
4. Choose **Add tag**.
5. In the **Key** box, enter the tag key for the tag to add to the resource. Then, in the **Value** box, optionally enter a tag value for the key.

   A tag key can contain as many as 128 characters. A tag value can contain as many as 256 characters. The characters can be letters, numbers, spaces, or the following symbols: _ . : / = + - @
6. To add another tag to the resource, choose **Add tag**, and then repeat the preceding step. You can assign as many as 50 tags to a resource.
7. When you finish adding tags, choose **Save**.

API

To create a resource and add one or more tags to it programmatically, use the appropriate `Create` operation for the type of resource that you want to create:

- **Allow list** – Use the CreateAllowList operation. Or, if you're using the AWS Command Line Interface (AWS CLI), run the create-allow-list command.
- **Custom data identifier** – Use the CreateCustomDataIdentifier operation. Or, if you're using the AWS CLI, run the create-custom-data-identifier command.
- **Filter or suppression rule** – Use the CreateFindingsFilter operation. Or, if you're using the AWS CLI, run the create-findings-filter command.
- **Member account** – Use the CreateMember operation. Or, if you're using the AWS CLI, run the create-member command.
- **Sensitive data discovery job** – Use the CreateClassificationJob operation. Or, if you're using the AWS CLI, run the create-classification-job command.

In your request, use the `tags` parameter to specify the tag key (`key`) and optional tag value (`value`) for each tag to add to the resource. The `tags` parameter specifies a string-to-string map of tag keys and their associated tag values.

To add one or more tags to an existing resource, use the TagResource operation of the Amazon Macie API or, if you're using the AWS CLI, run the tag-resource command. In your request, specify the Amazon Resource Name (ARN) of the resource that you want to add a tag to. Use the `tags` parameter to specify the tag key (`key`) and optional tag value (`value`) for each tag to add to the resource. As is the case for `Create` operations and commands, the `tags` parameter specifies a string-to-string map of tag keys and their associated tag values.

For example, the following AWS CLI command adds a `Stack` tag key with a `Production` tag value to the specified job. This example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws macie2 tag-resource ^
--resource-arn arn:aws:macie2:us-east-1:123456789012:classification-
job/3ce05dbb7ec5505def334104bexample ^
--tags={\"Stack\":\"Production\"}
```

Where:

- `resource-arn` specifies the ARN of the job to add a tag to.
- *`Stack`* is the tag key of the tag to add to the job.
- *`Production`* is the tag value for the specified tag key (*`Stack`*).

In the following example, the command adds several tags to the job:

```
C:\> aws macie2 tag-resource ^
--resource-arn arn:aws:macie2:us-east-1:123456789012:classification-job/3ce05dbb7ec5505def334104bexample ^
--tags={\"Stack\":\"Production\",\"CostCenter\":\"12345\",\"Owner\":\"jane-doe\"}
```

For each tag in a `tags` map, both the `key` and `value` arguments are required. However, the value for the `value` argument can be an empty string. If you don't want to associate a tag value with a tag key, don't specify a value for the `value` argument. For example, the following AWS CLI command adds an `Owner` tag key with no associated tag value:

```
C:\> aws macie2 tag-resource ^
--resource-arn arn:aws:macie2:us-east-1:123456789012:classification-job/3ce05dbb7ec5505def334104bexample ^
--tags={\"Owner\":\"\"}
```

If a tagging operation succeeds, Macie returns an empty HTTP 204 response. Otherwise, Macie returns an HTTP 4*xx* or 500 response that indicates why the operation failed.

Controlling access to Macie resources by using tags

After you start tagging Amazon Macie resources, you can define tag-based, resource-level permissions in AWS Identity and Access Management (IAM) policies. By using tags in this way, you can implement granular control of which users and roles in your AWS account have permission to create and tag Macie resources, and which users and roles have permission to add, edit, and remove tags more generally. To control access based on tags, you can use tag-related condition keys for Macie in the Condition element of IAM policies.

For example, you can create a policy that allows a user to have full access to all Macie resources, if the `Owner` tag for the resource specifies their username:

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
```

```
            "Sid": "ModifyResourceIfOwner",
            "Effect": "Allow",
            "Action": "macie2:*",
            "Resource": "*",
            "Condition": {
                "StringEqualsIgnoreCase": {"aws:ResourceTag/Owner":
 "${aws:username}"}
            }
        }
    ]
}
```

If you define tag-based, resource-level permissions, the permissions take effect immediately. This means that your resources are more secure as soon as they're created. It also means that you can quickly start enforcing the use of tags for new resources. You can also use resource-level permissions to control which tag keys and values can be associated with new and existing resources. For more information, see Controlling access to AWS resources using tags in the *IAM User Guide*.

Reviewing and editing tags for Macie resources

As your environment or requirements change over time, you can evaluate existing tags for your Amazon Macie resources and change the tags as necessary. A *tag* is a label that you define and assign to one or more AWS resources, including certain types of Macie resources. Each tag consists of a required *tag key* and an optional *tag value*. A *tag key* is a general label that acts as a category for a more specific tag value. A *tag value* acts as a descriptor for a tag key.

Tags can help you identify, categorize, and manage resources in different ways, such as by purpose, owner, environment, or other criteria. For example, you can use tags to: apply policies, allocate costs, distinguish between versions of resources, or identify resources that support certain compliance requirements or workflows.

You can assign tags to the following types of Macie resources:

- Allow lists
- Custom data identifiers
- Filter rules and suppression rules for findings
- Sensitive data discovery jobs

If you're the Macie administrator for an organization, you can also assign tags to member accounts in your organization. A resource can have as many as 50 tags.

Topics

Reviewing tags for Macie resources

You can review the tags for an Amazon Macie resource by using Macie or AWS Resource Groups. AWS Resource Groups is a service that's designed to help you group and manage AWS resources as a single unit instead of individually. If you use Macie, you can review the tags for one resource at a time. With AWS Resource Groups, you can review tags in bulk for multiple existing resources spanning multiple AWS services, including Macie.

To review the tags for a Macie resource

To review the tags for an individual Macie resource, you can use the Amazon Macie console or the Amazon Macie API. To review tags for multiple Macie resources at the same time, use the Tag Editor on the AWS Resource Groups console or the tagging operations of the AWS Resource Groups Tagging API. For more information, see the Tagging AWS Resources User Guide.

Console

Follow these steps to review a resource's tags by using the Amazon Macie console.

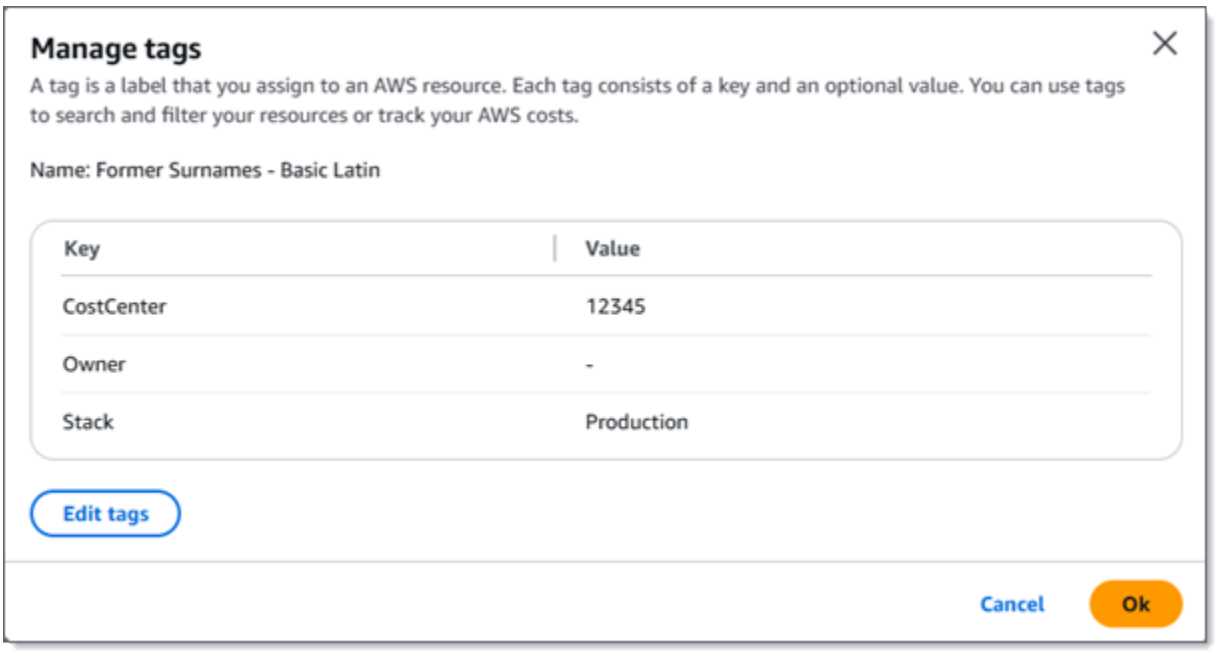
To review the tags for a resource

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. Depending on the type of resource whose tags you want to review, do one of the following:
   - For an allow list, choose **Allow lists** in the navigation pane. In the table, select the checkbox for the list. Then choose **Manage tags** on the **Actions** menu.
   - For a custom data identifier, choose **Custom data identifiers** in the navigation pane. In the table, select the checkbox for the custom data identifier. Then choose **Manage tags** on the **Actions** menu.
   - For a filter or suppression rule, choose **Findings** in the navigation pane. In the **Saved rules** list, choose the edit icon

( ) next to the rule. Then choose **Manage tags**.

- For a member account in your organization, choose **Accounts** in the navigation pane. In the table, select the checkbox for the account. Then choose **Manage tags** on the **Actions** menu.
- For a sensitive data discovery job, choose **Jobs** in the navigation pane. In the table, select the checkbox for the job. Then choose **Manage tags** on the **Actions** menu.

The **Manage tags** window lists all the tags that are currently assigned to the resource. For example, the following image shows the tags that are assigned to a custom data identifier.



In this example, three tags are assigned to the custom data identifier: the **CostCenter** tag key with **12345** as an associated tag value; the **Owner** tag key with no associated tag value (–); and, the **Stack** tag key with **Production** as an associated tag value.

3. When you finish reviewing the tags, choose **Cancel** to close the window.

API

To retrieve and review the tags for an existing resource programmatically, you can use the appropriate `Get` or `Describe` operation for the type of resource whose tags you want to review. For example, if you use the GetCustomDataIdentifier operation or you run the get-custom-data-identifier command from the AWS Command Line Interface (AWS CLI), the

response includes a `tags` object. The object lists all the tags (both tag keys and tag values) that are currently assigned to the resource.

You can also use the ListTagsForResource operation of the Amazon Macie API. In your request, use the `resourceArn` parameter to specify the Amazon Resource Name (ARN) of the resource. If you're using the AWS CLI, run the list-tags-for-resource command and use the `resource-arn` parameter to specify the ARN of the resource. For example:

```
C:\> aws macie2 list-tags-for-resource --resource-arn arn:aws:macie2:us-east-1:123456789012:classification-job/3ce05dbb7ec5505def334104bexample
```

In the preceding example, *arn:aws:macie2:us-east-1:123456789012:classification-job/3ce05dbb7ec5505def334104bexample* is the ARN of an existing sensitive data discovery job.

If the operation succeeds, Macie returns a `tags` object that lists all the tags (both tag keys and tag values) that are currently assigned to the resource. For example:

```
{
    "tags": {
        "Stack": "Production",
        "CostCenter": "12345",
        "Owner": ""
    }
}
```

Where `Stack`, `CostCenter`, and `Owner` are the tag keys that are assigned to the resource. `Production` is the tag value that's associated with the `Stack` tag key. `12345` is the tag value that's associated with the `CostCenter` tag key. The `Owner` tag key doesn't have an associated tag value.

To retrieve a list of all the Macie resources that have tags and all the tags that are assigned to each of those resources, use the GetResources operation of the AWS Resource Groups Tagging API. In your request, set the value for the `ResourceTypeFilters` parameter to `macie2`. To do this by using the AWS CLI, run the get-resources command and set the value for the `resource-type-filters` parameter to `macie2`. For example:

```
C:\> aws resourcegroupstaggingapi get-resources --resource-type-filters "macie2"
```

If the operation succeeds, Resource Groups returns a `ResourceTagMappingList` array that contains the ARNs of all the Macie resources that have tags, and the tag keys and values that are assigned to each of those resources.

Editing tags for Macie resources

To edit the tags (tag keys or tag values) for an Amazon Macie resource, you can use Macie or AWS Resource Groups. If you use Macie, you can edit the tags for one resource at a time. If you use AWS Resource Groups, you can edit tags in bulk for multiple existing resources spanning multiple AWS services, including Macie.

To edit the tags for a Macie resource

To edit the tags for an individual Macie resource, you can use the Amazon Macie console or the Amazon Macie API. To edit tags for multiple Macie resources at the same time, use the Tag Editor on the AWS Resource Groups console or the tagging operations of the AWS Resource Groups Tagging API.

> **Important**
>
> Editing the tags for a resource can affect access to the resource. Before you edit a tag key or value for a resource, review any AWS Identity and Access Management (IAM) policies that might use the tag to control access to resources. For more information, see Controlling access to AWS resources using tags in the *IAM User Guide*.

Console

Follow these steps to edit a resource's tags by using the Amazon Macie console.

To edit the tags for a resource

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. Depending on the type of resource whose tags you want to edit, do one of the following:
   - For an allow list, choose **Allow lists** in the navigation pane. In the table, select the checkbox for the list. Then choose **Manage tags** on the **Actions** menu.

- For a custom data identifier, choose **Custom data identifiers** in the navigation pane. In the table, select the checkbox for the custom data identifier. Then choose **Manage tags** on the **Actions** menu.
- For a filter or suppression rule, choose **Findings** in the navigation pane. In the **Saved rules** list, choose the edit icon () next to the rule. Then choose **Manage tags**.
- For a member account in your organization, choose **Accounts** in the navigation pane. In the table, select the checkbox for the account. Then choose **Manage tags** on the **Actions** menu.
- For a sensitive data discovery job, choose **Jobs** in the navigation pane. In the table, select the checkbox for the job. Then choose **Manage tags** on the **Actions** menu.

The **Manage tags** window lists all the tags that are currently assigned to the resource.

3. In the **Manage tags** window, choose **Edit tags**.
4. Do any of the following:

   - To add a tag value to a tag key, enter the value in the **Value** box next to the tag key.
   - To change an existing tag key, choose **Remove** next to the tag. Then choose **Add tag**. In the **Key** box that appears, enter the new tag key. Optionally enter an associated tag value in the **Value** box.
   - To change an existing tag value, choose **X** in the **Value** box that contains the value. Then enter the new tag value in the **Value** box.
   - To remove an existing tag value, choose **X** in the **Value** box that contains the value.
   - To remove an existing tag (both the tag key and tag value), choose **Remove** next to the tag.

   A resource can have as many as 50 tags. A tag key can contain as many as 128 characters. A tag value can contain as many as 256 characters. The characters can be letters, numbers, spaces, or the following symbols: _ . : / = + - @

5. When you finish editing the tags, choose **Save**.

API

When you edit a tag for a resource programmatically, you overwrite the existing tag with new values. Therefore, the best way to edit a tag depends on whether you want to edit a tag key, a tag value, or both. To edit a tag key, remove the current tag and add a new tag.

To edit or remove only the tag value that's associated with a tag key, overwrite the existing value by using the TagResource operation of the Amazon Macie API. If you're using the AWS Command Line Interface (AWS CLI), you can do this by running the tag-resource command. In your request, specify the Amazon Resource Name (ARN) of the resource whose tag value you want to edit or remove.

To edit a tag value for a tag key, use the `tags` parameter to specify the tag key whose tag value you want to change, and specify the new tag value for the key. For example, the following command changes the tag value from `Production` to `Staging` for the `Stack` tag key that's assigned to the specified sensitive data discovery job. This example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws macie2 tag-resource ^
--resource-arn arn:aws:macie2:us-east-1:123456789012:classification-
job/3ce05dbb7ec5505def334104bexample ^
--tags={\"Stack\":\"Staging\"}
```

Where:

- `resource-arn` specifies the job's ARN.
- *`Stack`* is the tag key that's associated with the tag value to change.
- *`Staging`* is the new tag value for the specified tag key (*`Stack`*).

To remove a tag value from a tag key, don't specify a value for the `value` argument in the `tags` parameter. For example:

```
C:\> aws macie2 tag-resource ^
--resource-arn arn:aws:macie2:us-east-1:123456789012:classification-
job/3ce05dbb7ec5505def334104bexample ^
--tags={\"Stack\":\"\"}
```

If the operation succeeds, Macie returns an empty HTTP 204 response. Otherwise, Macie returns an HTTP 4*xx* or 500 response that indicates why the operation failed.

Removing tags from Macie resources

If you add tags to an Amazon Macie resource, you can subsequently remove one or more of them. A *tag* is a label that you define and assign to AWS resources, including certain types of Macie resources. You can add, edit, and remove tags from the following types of Macie resources: allow lists, custom data identifiers, filter rules and suppression rules for findings, member accounts in an organization, and sensitive data discovery jobs.

You can remove tags from a Macie resource by using Macie or AWS Resource Groups. AWS Resource Groups is a service that's designed to help you group and manage AWS resources as a single unit instead of individually. If you use Macie, you can remove tags from one resource at a time. With AWS Resource Groups, you can remove tags in bulk for multiple existing resources spanning multiple AWS services, including Macie.

To remove tags from a Macie resource

To remove tags from a Macie resource, you can use the Amazon Macie console or the Amazon Macie API. To do this for multiple Macie resources at the same time, use the Tag Editor on the AWS Resource Groups console or the tagging operations of the AWS Resource Groups Tagging API. For more information, see the Tagging AWS Resources User Guide.

> **Important**
>
> Removing tags from a resource can affect access to the resource. Before you remove a tag, review any AWS Identity and Access Management (IAM) policies that might use the tag to control access to resources. For more information, see Controlling access to AWS resources using tags in the *IAM User Guide*.

Console

Follow these steps to remove one or more tags from a resource by using the Amazon Macie console.

To remove a tag from a resource

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. Depending on the type of resource that you want to remove a tag from, do one of the following:

- For an allow list, choose **Allow lists** in the navigation pane. In the table, select the checkbox for the list. Then choose **Manage tags** on the **Actions** menu.
- For a custom data identifier, choose **Custom data identifiers** in the navigation pane. In the table, select the checkbox for the custom data identifier. Then choose **Manage tags** on the **Actions** menu.
- For a filter or suppression rule, choose **Findings** in the navigation pane. In the **Saved rules** list, choose the edit icon () next to the rule. Then choose **Manage tags**.
- For a member account in your organization, choose **Accounts** in the navigation pane. In the table, select the checkbox for the account. Then choose **Manage tags** on the **Actions** menu.
- For a sensitive data discovery job, choose **Jobs** in the navigation pane. In the table, select the checkbox for the job. Then choose **Manage tags** on the **Actions** menu.

The **Manage tags** window lists all the tags that are currently assigned to the resource.

3. In the **Manage tags** window, choose **Edit tags**.
4. Do any of the following:
   - To remove only the tag value for a tag, choose **X** in the **Value** box that contains the value to remove.
   - To remove both the tag key and tag value (as a pair) for a tag, choose **Remove** next to the tag to remove.
5. To remove additional tags from the resource, repeat the preceding step for each additional tag to remove.
6. When you finish removing tags, choose **Save**.

API

To remove one or more tags from a resource programmatically, use the UntagResource operation of the Amazon Macie API. In your request, use the `resourceArn` parameter to specify the Amazon Resource Name (ARN) of the resource to remove a tag from. Use the `tagKeys` parameter to specify the tag key of the tag to remove. To remove only a specific tag value (not a tag key) from a resource, edit the tag instead of removing the tag.

If you're using the AWS Command Line Interface (AWS CLI), run the untag-resource command and use the `resource-arn` parameter to specify the ARN of the resource to remove a tag from. Use the `tag-keys` parameter to specify the tag key of the tag to remove. For example, the following command removes the `Stack` tag (both the tag key and tag value) from the specified sensitive data discovery job:

```
C:\> aws macie2 untag-resource ^
--resource-arn arn:aws:macie2:us-east-1:123456789012:classification-
job/3ce05dbb7ec5505def334104bexample ^
--tag-keys Stack
```

Where `resource-arn` specifies the ARN of the job to remove a tag from, and *`Stack`* is the tag key of the tag to remove.

To remove multiple tags from a resource, add each additional tag key as an argument for the `tag-keys` parameter. For example:

```
C:\> aws macie2 untag-resource ^
--resource-arn arn:aws:macie2:us-east-1:123456789012:classification-
job/3ce05dbb7ec5505def334104bexample ^
--tag-keys Stack Owner
```

Where `resource-arn` specifies the ARN of the job to remove tags from, and *`Stack`* and *`Owner`* are the tag keys of the tags to remove.

If the operation succeeds, Macie returns an empty HTTP 204 response. Otherwise, Macie returns an HTTP 4*xx* or 500 response that indicates why the operation failed.

Security in Macie

Cloud security at AWS is the highest priority. As an AWS customer, you benefit from data centers and network architectures that are built to meet the requirements of the most security-sensitive organizations.

Security is a shared responsibility between AWS and you. The shared responsibility model describes this as security *of* the cloud and security *in* the cloud:

- **Security of the cloud** – AWS is responsible for protecting the infrastructure that runs AWS services in the AWS Cloud. AWS also provides you with services that you can use securely. Third-party auditors regularly test and verify the effectiveness of our security as part of the AWS Compliance Programs. To learn about the compliance programs that apply to Amazon Macie, see AWS Services in Scope by Compliance Program.
- **Security in the cloud** – Your responsibility is determined by the AWS services that you use. You are also responsible for other factors including the sensitivity of your data, your company's requirements, and applicable laws and regulations.

This documentation helps you understand how to apply the shared responsibility model when using Amazon Macie. The following topics show you how to configure Macie to meet your security and compliance objectives. You also learn how to use other AWS services that can help you monitor and secure your Macie resources.

Topics

Data protection in Macie

The AWS shared responsibility model applies to data protection in Amazon Macie. As described in this model, AWS is responsible for protecting the global infrastructure that runs all of the

AWS Cloud. You are responsible for maintaining control over your content that is hosted on this infrastructure. You are also responsible for the security configuration and management tasks for the AWS services that you use. For more information about data privacy, see the Data Privacy FAQ. For information about data protection in Europe, see the AWS Shared Responsibility Model and GDPR blog post on the *AWS Security Blog*.

For data protection purposes, we recommend that you protect AWS account credentials and set up individual users with AWS IAM Identity Center or AWS Identity and Access Management (IAM). That way, each user is given only the permissions necessary to fulfill their job duties. We also recommend that you secure your data in the following ways:

- Use multi-factor authentication (MFA) with each account.
- Use SSL/TLS to communicate with AWS resources. We require TLS 1.2 and recommend TLS 1.3.
- Set up API and user activity logging with AWS CloudTrail. For information about using CloudTrail trails to capture AWS activities, see Working with CloudTrail trails in the *AWS CloudTrail User Guide*.
- Use AWS encryption solutions, along with all default security controls within AWS services.
- Use advanced managed security services such as Amazon Macie, which assists in discovering and securing sensitive data that is stored in Amazon S3.
- If you require FIPS 140-3 validated cryptographic modules when accessing AWS through a command line interface or an API, use a FIPS endpoint. For more information about the available FIPS endpoints, see Federal Information Processing Standard (FIPS) 140-3.

We strongly recommend that you never put confidential or sensitive information, such as your customers' email addresses, into tags or free-form text fields such as a **Name** field. This includes when you work with Macie or other AWS services using the console, API, AWS CLI, or AWS SDKs. Any data that you enter into tags or free-form text fields used for names may be used for billing or diagnostic logs. If you provide a URL to an external server, we strongly recommend that you do not include credentials information in the URL to validate your request to that server.

Encryption at rest

Amazon Macie securely stores your data at rest using AWS encryption solutions. Macie encrypts data, such as findings, using an AWS managed key from AWS Key Management Service (AWS KMS).

If you disable Macie, it permanently deletes all resources that it stores or maintains for you, such as sensitive data discovery jobs, custom data identifiers, and findings.

Encryption in transit

Amazon Macie encrypts all data in transit between AWS services.

Macie analyzes data from Amazon S3 and exports sensitive data discovery results to an S3 general purpose bucket. After Macie gets the information that it needs from S3 objects, the objects are discarded.

Macie accesses Amazon S3 by using a VPC endpoint powered by AWS PrivateLink. Therefore, traffic between Macie and Amazon S3 stays on the Amazon network and does not go over the public internet. For more information, see AWS PrivateLink.

Identity and access management for Macie

AWS Identity and Access Management (IAM) is an AWS service that helps an administrator securely control access to AWS resources. IAM administrators control who can be *authenticated* (signed in) and *authorized* (have permissions) to use Macie resources. IAM is an AWS service that you can use with no additional charge.

Topics

Audience

How you use AWS Identity and Access Management (IAM) differs, depending on the work that you do in Macie.

Service user – If you use the Macie service to do your job, then your administrator provides you with the credentials and permissions that you need. As you use more Macie features to do your

work, you might need additional permissions. Understanding how access is managed can help you request the right permissions from your administrator. If you cannot access a feature in Macie, see Troubleshooting identity and access management for Macie.

Service administrator – If you're in charge of Macie resources at your company, you probably have full access to Macie. It's your job to determine which Macie features and resources your service users should access. You must then submit requests to your IAM administrator to change the permissions of your service users. Review the information on this page to understand the basic concepts of IAM. To learn more about how your company can use IAM with Macie, see How Macie works with AWS Identity and Access Management.

IAM administrator – If you're an IAM administrator, you might want to learn details about how you can write policies to manage access to Macie. To view example Macie identity-based policies that you can use in IAM, see Identity-based policy examples for Macie.

Authenticating with identities

Authentication is how you sign in to AWS using your identity credentials. You must be *authenticated* (signed in to AWS) as the AWS account root user, as an IAM user, or by assuming an IAM role.

You can sign in to AWS as a federated identity by using credentials provided through an identity source. AWS IAM Identity Center (IAM Identity Center) users, your company's single sign-on authentication, and your Google or Facebook credentials are examples of federated identities. When you sign in as a federated identity, your administrator previously set up identity federation using IAM roles. When you access AWS by using federation, you are indirectly assuming a role.

Depending on the type of user you are, you can sign in to the AWS Management Console or the AWS access portal. For more information about signing in to AWS, see How to sign in to your AWS account in the *AWS Sign-In User Guide*.

If you access AWS programmatically, AWS provides a software development kit (SDK) and a command line interface (CLI) to cryptographically sign your requests by using your credentials. If you don't use AWS tools, you must sign requests yourself. For more information about using the recommended method to sign requests yourself, see AWS Signature Version 4 for API requests in the *IAM User Guide*.

Regardless of the authentication method that you use, you might be required to provide additional security information. For example, AWS recommends that you use multi-factor authentication (MFA) to increase the security of your account. To learn more, see Multi-factor authentication in

the *AWS IAM Identity Center User Guide* and AWS Multi-factor authentication in IAM in the *IAM User Guide*.

AWS account root user

When you create an AWS account, you begin with one sign-in identity that has complete access to all AWS services and resources in the account. This identity is called the AWS account *root user* and is accessed by signing in with the email address and password that you used to create the account. We strongly recommend that you don't use the root user for your everyday tasks. Safeguard your root user credentials and use them to perform the tasks that only the root user can perform. For the complete list of tasks that require you to sign in as the root user, see Tasks that require root user credentials in the *IAM User Guide*.

Federated identity

As a best practice, require human users, including users that require administrator access, to use federation with an identity provider to access AWS services by using temporary credentials.

A *federated identity* is a user from your enterprise user directory, a web identity provider, the AWS Directory Service, the Identity Center directory, or any user that accesses AWS services by using credentials provided through an identity source. When federated identities access AWS accounts, they assume roles, and the roles provide temporary credentials.

For centralized access management, we recommend that you use AWS IAM Identity Center. You can create users and groups in IAM Identity Center, or you can connect and synchronize to a set of users and groups in your own identity source for use across all your AWS accounts and applications. For information about IAM Identity Center, see What is IAM Identity Center? in the *AWS IAM Identity Center User Guide*.

IAM users and groups

An *IAM user* is an identity within your AWS account that has specific permissions for a single person or application. Where possible, we recommend relying on temporary credentials instead of creating IAM users who have long-term credentials such as passwords and access keys. However, if you have specific use cases that require long-term credentials with IAM users, we recommend that you rotate access keys. For more information, see Rotate access keys regularly for use cases that require long-term credentials in the *IAM User Guide*.

An *IAM group* is an identity that specifies a collection of IAM users. You can't sign in as a group. You can use groups to specify permissions for multiple users at a time. Groups make permissions easier

to manage for large sets of users. For example, you could have a group named *IAMAdmins* and give that group permissions to administer IAM resources.

Users are different from roles. A user is uniquely associated with one person or application, but a role is intended to be assumable by anyone who needs it. Users have permanent long-term credentials, but roles provide temporary credentials. To learn more, see Use cases for IAM users in the *IAM User Guide*.

IAM roles

An *IAM role* is an identity within your AWS account that has specific permissions. It is similar to an IAM user, but is not associated with a specific person. To temporarily assume an IAM role in the AWS Management Console, you can switch from a user to an IAM role (console). You can assume a role by calling an AWS CLI or AWS API operation or by using a custom URL. For more information about methods for using roles, see Methods to assume a role in the *IAM User Guide*.

IAM roles with temporary credentials are useful in the following situations:

- **Federated user access** – To assign permissions to a federated identity, you create a role and define permissions for the role. When a federated identity authenticates, the identity is associated with the role and is granted the permissions that are defined by the role. For information about roles for federation, see Create a role for a third-party identity provider (federation) in the *IAM User Guide*. If you use IAM Identity Center, you configure a permission set. To control what your identities can access after they authenticate, IAM Identity Center correlates the permission set to a role in IAM. For information about permissions sets, see Permission sets in the *AWS IAM Identity Center User Guide*.
- **Temporary IAM user permissions** – An IAM user or role can assume an IAM role to temporarily take on different permissions for a specific task.
- **Cross-account access** – You can use an IAM role to allow someone (a trusted principal) in a different account to access resources in your account. Roles are the primary way to grant cross-account access. However, with some AWS services, you can attach a policy directly to a resource (instead of using a role as a proxy). To learn the difference between roles and resource-based policies for cross-account access, see Cross account resource access in IAM in the *IAM User Guide*.
- **Cross-service access** – Some AWS services use features in other AWS services. For example, when you make a call in a service, it's common for that service to run applications in Amazon EC2 or store objects in Amazon S3. A service might do this using the calling principal's permissions, using a service role, or using a service-linked role.

- **Forward access sessions (FAS)** – When you use an IAM user or role to perform actions in AWS, you are considered a principal. When you use some services, you might perform an action that then initiates another action in a different service. FAS uses the permissions of the principal calling an AWS service, combined with the requesting AWS service to make requests to downstream services. FAS requests are only made when a service receives a request that requires interactions with other AWS services or resources to complete. In this case, you must have permissions to perform both actions. For policy details when making FAS requests, see Forward access sessions.
- **Service role** – A service role is an IAM role that a service assumes to perform actions on your behalf. An IAM administrator can create, modify, and delete a service role from within IAM. For more information, see Create a role to delegate permissions to an AWS service in the *IAM User Guide*.
- **Service-linked role** – A service-linked role is a type of service role that is linked to an AWS service. The service can assume the role to perform an action on your behalf. Service-linked roles appear in your AWS account and are owned by the service. An IAM administrator can view, but not edit the permissions for service-linked roles.

- **Applications running on Amazon EC2** – You can use an IAM role to manage temporary credentials for applications that are running on an EC2 instance and making AWS CLI or AWS API requests. This is preferable to storing access keys within the EC2 instance. To assign an AWS role to an EC2 instance and make it available to all of its applications, you create an instance profile that is attached to the instance. An instance profile contains the role and enables programs that are running on the EC2 instance to get temporary credentials. For more information, see Use an IAM role to grant permissions to applications running on Amazon EC2 instances in the *IAM User Guide*.

Managing access using policies

You control access in AWS by creating policies and attaching them to AWS identities or resources. A policy is an object in AWS that, when associated with an identity or resource, defines their permissions. AWS evaluates these policies when a principal (user, root user, or role session) makes a request. Permissions in the policies determine whether the request is allowed or denied. Most policies are stored in AWS as JSON documents. For more information about the structure and contents of JSON policy documents, see Overview of JSON policies in the *IAM User Guide*.

Administrators can use AWS JSON policies to specify who has access to what. That is, which **principal** can perform **actions** on what **resources**, and under what **conditions**.

By default, users and roles have no permissions. To grant users permission to perform actions on the resources that they need, an IAM administrator can create IAM policies. The administrator can then add the IAM policies to roles, and users can assume the roles.

IAM policies define permissions for an action regardless of the method that you use to perform the operation. For example, suppose that you have a policy that allows the `iam:GetRole` action. A user with that policy can get role information from the AWS Management Console, the AWS CLI, or the AWS API.

Identity-based policies

Identity-based policies are JSON permissions policy documents that you can attach to an identity, such as an IAM user, group of users, or role. These policies control what actions users and roles can perform, on which resources, and under what conditions. To learn how to create an identity-based policy, see Define custom IAM permissions with customer managed policies in the *IAM User Guide*.

Identity-based policies can be further categorized as *inline policies* or *managed policies*. Inline policies are embedded directly into a single user, group, or role. Managed policies are standalone policies that you can attach to multiple users, groups, and roles in your AWS account. Managed policies include AWS managed policies and customer managed policies. To learn how to choose between a managed policy or an inline policy, see Choose between managed policies and inline policies in the *IAM User Guide*.

Resource-based policies

Resource-based policies are JSON policy documents that you attach to a resource. Examples of resource-based policies are IAM *role trust policies* and Amazon S3 *bucket policies*. In services that support resource-based policies, service administrators can use them to control access to a specific resource. For the resource where the policy is attached, the policy defines what actions a specified principal can perform on that resource and under what conditions. You must specify a principal in a resource-based policy. Principals can include accounts, users, roles, federated users, or AWS services.

Resource-based policies are inline policies that are located in that service. You can't use AWS managed policies from IAM in a resource-based policy.

Access control lists (ACLs)

Access control lists (ACLs) control which principals (account members, users, or roles) have permissions to access a resource. ACLs are similar to resource-based policies, although they do not use the JSON policy document format.

Amazon S3, AWS WAF, and Amazon VPC are examples of services that support ACLs. To learn more about ACLs, see Access control list (ACL) overview in the *Amazon Simple Storage Service Developer Guide*.

Other policy types

AWS supports additional, less-common policy types. These policy types can set the maximum permissions granted to you by the more common policy types.

- **Permissions boundaries** – A permissions boundary is an advanced feature in which you set the maximum permissions that an identity-based policy can grant to an IAM entity (IAM user or role). You can set a permissions boundary for an entity. The resulting permissions are the intersection of an entity's identity-based policies and its permissions boundaries. Resource-based policies that specify the user or role in the `Principal` field are not limited by the permissions boundary. An explicit deny in any of these policies overrides the allow. For more information about permissions boundaries, see Permissions boundaries for IAM entities in the *IAM User Guide*.
- **Service control policies (SCPs)** – SCPs are JSON policies that specify the maximum permissions for an organization or organizational unit (OU) in AWS Organizations. AWS Organizations is a service for grouping and centrally managing multiple AWS accounts that your business owns. If you enable all features in an organization, then you can apply service control policies (SCPs) to any or all of your accounts. The SCP limits permissions for entities in member accounts, including each AWS account root user. For more information about Organizations and SCPs, see Service control policies in the *AWS Organizations User Guide*.
- **Resource control policies (RCPs)** – RCPs are JSON policies that you can use to set the maximum available permissions for resources in your accounts without updating the IAM policies attached to each resource that you own. The RCP limits permissions for resources in member accounts and can impact the effective permissions for identities, including the AWS account root user, regardless of whether they belong to your organization. For more information about Organizations and RCPs, including a list of AWS services that support RCPs, see Resource control policies (RCPs) in the *AWS Organizations User Guide*.
- **Session policies** – Session policies are advanced policies that you pass as a parameter when you programmatically create a temporary session for a role or federated user. The resulting session's

permissions are the intersection of the user or role's identity-based policies and the session policies. Permissions can also come from a resource-based policy. An explicit deny in any of these policies overrides the allow. For more information, see Session policies in the *IAM User Guide*.

Multiple policy types

When multiple types of policies apply to a request, the resulting permissions are more complicated to understand. To learn how AWS determines whether to allow a request when multiple policy types are involved, see Policy evaluation logic in the *IAM User Guide*.

How Macie works with AWS Identity and Access Management

Before you use AWS Identity and Access Management (IAM) to manage access to Amazon Macie, learn which IAM features are available to use with Macie.

IAM features you can use with Macie

IAM feature	Macie support
Identity-based policies	Yes
Resource-based policies	No
Policy actions	Yes
Policy resources	Yes
Policy condition keys	Yes
Access control lists (ACLs)	No
Attribute-based access control (ABAC) – tags in policies	Yes
Temporary credentials	Yes
Forward access sessions (FAS)	Yes
Service roles	No

IAM feature	Macie support
Service-linked roles	Yes

For a high-level view of how Macie and other AWS services work with most IAM features, see AWS services that work with IAM in the *IAM User Guide*.

Identity-based policies for Macie

Supports identity-based policies: Yes

Identity-based policies are JSON permissions policy documents that you can attach to an identity, such as an IAM user, group of users, or role. These policies control what actions users and roles can perform, on which resources, and under what conditions. To learn how to create an identity-based policy, see Define custom IAM permissions with customer managed policies in the *IAM User Guide*.

With IAM identity-based policies, you can specify allowed or denied actions and resources as well as the conditions under which actions are allowed or denied. You can't specify the principal in an identity-based policy because it applies to the user or role to which it is attached. To learn about all of the elements that you can use in a JSON policy, see IAM JSON policy elements reference in the *IAM User Guide*.

Amazon Macie supports identity-based policies. For examples, see Identity-based policy examples for Macie.

Resource-based policies within Macie

Supports resource-based policies: No

Resource-based policies are JSON policy documents that you attach to a resource. Examples of resource-based policies are IAM *role trust policies* and Amazon S3 *bucket policies*. In services that support resource-based policies, service administrators can use them to control access to a specific resource. For the resource where the policy is attached, the policy defines what actions a specified principal can perform on that resource and under what conditions. You must specify a principal in a resource-based policy. Principals can include accounts, users, roles, federated users, or AWS services.

To enable cross-account access, you can specify an entire account or IAM entities in another account as the principal in a resource-based policy. Adding a cross-account principal to a resource-

based policy is only half of establishing the trust relationship. When the principal and the resource are in different AWS accounts, an IAM administrator in the trusted account must also grant the principal entity (user or role) permission to access the resource. They grant permission by attaching an identity-based policy to the entity. However, if a resource-based policy grants access to a principal in the same account, no additional identity-based policy is required. For more information, see Cross account resource access in IAM in the *IAM User Guide*.

Amazon Macie doesn't support resource-based policies. That is to say, you can't attach a policy directly to a Macie resource.

Policy actions for Macie

Supports policy actions: Yes

Administrators can use AWS JSON policies to specify who has access to what. That is, which **principal** can perform **actions** on what **resources**, and under what **conditions**.

The `Action` element of a JSON policy describes the actions that you can use to allow or deny access in a policy. Policy actions usually have the same name as the associated AWS API operation. There are some exceptions, such as *permission-only actions* that don't have a matching API operation. There are also some operations that require multiple actions in a policy. These additional actions are called *dependent actions*.

Include actions in a policy to grant permissions to perform the associated operation.

Policy actions for Amazon Macie use the following prefix before the action:

```
macie2
```

For example, to grant someone permission to access information about all the managed data identifiers that Macie provides, which is an action that corresponds to the `ListManagedDataIdentifiers` operation of the Amazon Macie API, include the `macie2:ListManagedDataIdentifiers` action in their policy:

```
"Action": "macie2:ListManagedDataIdentifiers"
```

To specify multiple actions in a single statement, separate them with commas. For example:

```
"Action": [
      "macie2:ListManagedDataIdentifiers",
      "macie2:ListCustomDataIdentifiers"
```

```
]
```

You can also specify multiple actions by using wildcards (*). For example, to specify all actions that begin with the word `List`, include the following action:

```
"Action": "macie2:List*"
```

However, as a best practice, you should create policies that follow the principle of least privilege. In other words, you should create policies that include only the permissions that are required to perform a specific task.

For a list of Macie actions, see Actions defined by Amazon Macie in the *Service Authorization Reference*. For examples of policies that specify Macie actions, see Identity-based policy examples for Macie.

Policy resources for Macie

Supports policy resources: Yes

Administrators can use AWS JSON policies to specify who has access to what. That is, which **principal** can perform **actions** on what **resources**, and under what **conditions**.

The `Resource` JSON policy element specifies the object or objects to which the action applies. Statements must include either a `Resource` or a `NotResource` element. As a best practice, specify a resource using its Amazon Resource Name (ARN). You can do this for actions that support a specific resource type, known as *resource-level permissions*.

For actions that don't support resource-level permissions, such as listing operations, use a wildcard (*) to indicate that the statement applies to all resources.

```
"Resource": "*"
```

Amazon Macie defines the following resource types:

- Allow list
- Custom data identifier
- Filter or suppression rule, also referred to as a *findings filter*
- Member account

- Sensitive data discovery job, also referred to as a *classification job*

You can specify these types of resources in policies by using ARNs.

For example, to create a policy for the sensitive data discovery job that has the job ID *3ce05dbb7ec5505def334104bexample*, you can use the following ARN:

```
"Resource": "arn:aws:macie2:*:*:classification-job/3ce05dbb7ec5505def334104bexample"
```

Or, to specify all the sensitive data discovery jobs for a certain account, use a wildcard (*):

```
"Resource": "arn:aws:macie2:*:123456789012:classification-job/*"
```

Where *123456789012* is the account ID for the AWS account that created the jobs. As a best practice, however, you should create policies that follow the principle of least privilege. In other words, you should create policies that include only the permissions that are required to perform a specific task on a specific resource.

Some Macie actions can apply to multiple resources. For example, the `macie2:BatchGetCustomDataIdentifiers` action can retrieve the details of multiple custom data identifiers. In these cases, a principal must have permissions to access all the resources that the action applies to. To specify multiple resources in a single statement, separate the ARNs with commas:

```
"Resource": [
 "arn:aws:macie2:*:*:custom-data-identifier/12g4aff9-8e22-4f2b-b3fd-3063eexample",
 "arn:aws:macie2:*:*:custom-data-identifier/2d12c96a-8e78-4ca6-b1dc-8fd65example",
 "arn:aws:macie2:*:*:custom-data-identifier/4383a69d-4a1e-4a07-8715-208ddexample"
]
```

For a list of Macie resource types and the ARN syntax for each one, see Resource types defined by Amazon Macie in the *Service Authorization Reference*. To learn which actions you can specify with each resource type, see Actions defined by Amazon Macie in the *Service Authorization Reference*. For examples of policies that specify resources, see Identity-based policy examples for Macie.

Policy condition keys for Macie

Supports service-specific policy condition keys: Yes

Administrators can use AWS JSON policies to specify who has access to what. That is, which **principal** can perform **actions** on what **resources**, and under what **conditions**.

The `Condition` element (or `Condition` *block*) lets you specify conditions in which a statement is in effect. The `Condition` element is optional. You can create conditional expressions that use condition operators, such as equals or less than, to match the condition in the policy with values in the request.

If you specify multiple `Condition` elements in a statement, or multiple keys in a single `Condition` element, AWS evaluates them using a logical `AND` operation. If you specify multiple values for a single condition key, AWS evaluates the condition using a logical `OR` operation. All of the conditions must be met before the statement's permissions are granted.

You can also use placeholder variables when you specify conditions. For example, you can grant an IAM user permission to access a resource only if it is tagged with their IAM user name. For more information, see IAM policy elements: variables and tags in the *IAM User Guide*.

AWS supports global condition keys and service-specific condition keys. To see all AWS global condition keys, see AWS global condition context keys in the *IAM User Guide*.

For a list of Amazon Macie condition keys, see Condition keys for Amazon Macie in the *Service Authorization Reference*. To learn which actions and resources you can use a condition key with, see Actions defined by Amazon Macie. For examples of policies that use condition keys, see Identity-based policy examples for Macie.

Access control lists (ACLs) in Macie

Supports ACLs: No

Access control lists (ACLs) control which principals (account members, users, or roles) have permissions to access a resource. ACLs are similar to resource-based policies, although they do not use the JSON policy document format.

Amazon Simple Storage Service (Amazon S3) is an example of an AWS service that supports ACLs. To learn more, see Access control list (ACL) overview in the *Amazon Simple Storage Service User Guide*.

Amazon Macie doesn't support ACLs. That is to say, you can't attach an ACL to a Macie resource.

Attribute-based access control (ABAC) with Macie

Supports ABAC (tags in policies): Yes

Attribute-based access control (ABAC) is an authorization strategy that defines permissions based on attributes. In AWS, these attributes are called *tags*. You can attach tags to IAM entities (users or roles) and to many AWS resources. Tagging entities and resources is the first step of ABAC. Then you design ABAC policies to allow operations when the principal's tag matches the tag on the resource that they are trying to access.

ABAC is helpful in environments that are growing rapidly and helps with situations where policy management becomes cumbersome.

To control access based on tags, you provide tag information in the condition element of a policy using the `aws:ResourceTag/key-name`, `aws:RequestTag/key-name`, or `aws:TagKeys` condition keys.

If a service supports all three condition keys for every resource type, then the value is **Yes** for the service. If a service supports all three condition keys for only some resource types, then the value is **Partial**.

For more information about ABAC, see Define permissions with ABAC authorization in the *IAM User Guide*. To view a tutorial with steps for setting up ABAC, see Use attribute-based access control (ABAC) in the *IAM User Guide*.

You can attach tags to Amazon Macie resources—allow lists, custom data identifiers, filter rules and suppression rules, member accounts, and sensitive data discovery jobs. You can also control access to these types of resources by providing tag information in the `Condition` element of a policy. For information about attaching tags to resources, see Tagging Macie resources. For an example of an identity-based policy that controls access to a resource based on tags, see Identity-based policy examples for Macie.

Using temporary credentials with Macie

Supports temporary credentials: Yes

Some AWS services don't work when you sign in using temporary credentials. For additional information, including which AWS services work with temporary credentials, see AWS services that work with IAM in the *IAM User Guide*.

You are using temporary credentials if you sign in to the AWS Management Console using any method except a user name and password. For example, when you access AWS using your company's single sign-on (SSO) link, that process automatically creates temporary credentials. You also automatically create temporary credentials when you sign in to the console as a user and then

switch roles. For more information about switching roles, see Switch from a user to an IAM role (console) in the *IAM User Guide*.

You can manually create temporary credentials using the AWS CLI or AWS API. You can then use those temporary credentials to access AWS. AWS recommends that you dynamically generate temporary credentials instead of using long-term access keys. For more information, see Temporary security credentials in IAM.

Amazon Macie supports the use of temporary credentials.

Forward access sessions for Macie

Supports forward access sessions (FAS): Yes

When you use an IAM user or role to perform actions in AWS, you are considered a principal. When you use some services, you might perform an action that then initiates another action in a different service. FAS uses the permissions of the principal calling an AWS service, combined with the requesting AWS service to make requests to downstream services. FAS requests are only made when a service receives a request that requires interactions with other AWS services or resources to complete. In this case, you must have permissions to perform both actions. For policy details when making FAS requests, see Forward access sessions.

Amazon Macie makes FAS requests to downstream AWS services when you perform the following tasks:

- Create or update Macie settings for an allow list that's stored in an S3 bucket.
- Check the status of an allow list that's stored in an S3 bucket.
- Retrieve sensitive data samples from an affected S3 object by using IAM user credentials.
- Encrypt sensitive data samples that are retrieved using IAM user credentials or an IAM role.
- Enable Macie to integrate with AWS Organizations.
- Designate the delegated Macie administrator account for an organization in AWS Organizations.

For other tasks, Macie uses a service-linked role to perform actions on your behalf. For details about this role, see Using service-linked roles for Macie.

Service roles for Macie

Supports service roles: No

A service role is an IAM role that a service assumes to perform actions on your behalf. An IAM administrator can create, modify, and delete a service role from within IAM. For more information, see Create a role to delegate permissions to an AWS service in the *IAM User Guide*.

Amazon Macie doesn't assume or use service roles. To perform actions on your behalf, Macie primarily uses a service-linked role. For details about this role, see Using service-linked roles for Macie.

Service-linked roles for Macie

Supports service-linked roles: Yes

A service-linked role is a type of service role that is linked to an AWS service. The service can assume the role to perform an action on your behalf. Service-linked roles appear in your AWS account and are owned by the service. An IAM administrator can view, but not edit the permissions for service-linked roles.

Amazon Macie uses a service-linked role to perform actions on your behalf. For details about this role, see Using service-linked roles for Macie.

Identity-based policy examples for Macie

By default, users and roles don't have permission to create or modify Macie resources. They also can't perform tasks by using the AWS Management Console, AWS Command Line Interface (AWS CLI), or AWS API. To grant users permission to perform actions on the resources that they need, an IAM administrator can create IAM policies. The administrator can then add the IAM policies to roles, and users can assume the roles.

To learn how to create an IAM identity-based policy by using these example JSON policy documents, see Create IAM policies (console) in the *IAM User Guide*.

For details about actions and resource types defined by Macie, including the format of the ARNs for each of the resource types, see Actions, resources, and condition keys for Amazon Macie in the *Service Authorization Reference*.

When you create a policy, be sure to resolve security warnings, errors, general warnings, and suggestions from AWS Identity and Access Management Access Analyzer (IAM Access Analyzer) before you save the policy. IAM Access Analyzer runs policy checks to validate a policy against IAM policy grammar and best practices. These checks generate findings and provide actionable recommendations to help you author policies that are functional and conform to security best practices. To learn about validating policies by using IAM Access Analyzer, see IAM Access Analyzer

policy validation in the *IAM User Guide*. To review a list of the warnings, errors, and suggestions that IAM Access Analyzer can return, see IAM Access Analyzer policy check reference in the *IAM User Guide*.

Topics

Policy best practices

Identity-based policies determine whether someone can create, access, or delete Macie resources in your account. These actions can incur costs for your AWS account. When you create or edit identity-based policies, follow these guidelines and recommendations:

- **Get started with AWS managed policies and move toward least-privilege permissions** – To get started granting permissions to your users and workloads, use the *AWS managed policies* that grant permissions for many common use cases. They are available in your AWS account. We recommend that you reduce permissions further by defining AWS customer managed policies that are specific to your use cases. For more information, see AWS managed policies or AWS managed policies for job functions in the *IAM User Guide*.
- **Apply least-privilege permissions** – When you set permissions with IAM policies, grant only the permissions required to perform a task. You do this by defining the actions that can be taken on specific resources under specific conditions, also known as *least-privilege permissions*. For more information about using IAM to apply permissions, see Policies and permissions in IAM in the *IAM User Guide*.
- **Use conditions in IAM policies to further restrict access** – You can add a condition to your policies to limit access to actions and resources. For example, you can write a policy condition to specify that all requests must be sent using SSL. You can also use conditions to grant access to service actions if they are used through a specific AWS service, such as AWS CloudFormation. For more information, see IAM JSON policy elements: Condition in the *IAM User Guide*.

- **Use IAM Access Analyzer to validate your IAM policies to ensure secure and functional permissions** – IAM Access Analyzer validates new and existing policies so that the policies adhere to the IAM policy language (JSON) and IAM best practices. IAM Access Analyzer provides more than 100 policy checks and actionable recommendations to help you author secure and functional policies. For more information, see Validate policies with IAM Access Analyzer in the *IAM User Guide*.
- **Require multi-factor authentication (MFA)** – If you have a scenario that requires IAM users or a root user in your AWS account, turn on MFA for additional security. To require MFA when API operations are called, add MFA conditions to your policies. For more information, see Secure API access with MFA in the *IAM User Guide*.

For more information about best practices in IAM, see Security best practices in IAM in the *IAM User Guide*.

Using the Amazon Macie console

To access the Amazon Macie console, you must have a minimum set of permissions. These permissions must allow you to list and view details about the Macie resources in your AWS account. If you create an identity-based policy that is more restrictive than the minimum required permissions, the console won't function as intended for entities (users or roles) with that policy.

You don't need to allow minimum console permissions for users that are making calls only to the AWS CLI or the AWS API. Instead, allow access to only the actions that match the API operation that they're trying to perform.

To ensure that users and roles can use the Amazon Macie console, create IAM policies that provide them with console access. For more information, see Policies and permissions in IAM in the *IAM User Guide*.

If you create a policy that allows users or roles to use the Amazon Macie console, ensure that the policy allows the `macie2:GetMacieSession` action. Otherwise, those users or roles won't be able to access any Macie resources or data on the console.

Also ensure that the policy allows the appropriate `macie2:List` actions for resources that those users or roles need to access on the console. Otherwise, they won't be able to navigate to or display details about those resources on the console. For example, to review the details of a sensitive data discovery job by using the console, a user must be allowed to perform the `macie2:DescribeClassificationJob` action for the job *and*

the `macie2:ListClassificationJobs` action. If a user isn't allowed to perform the `macie2:ListClassificationJobs` action, the user won't be able to display a list of jobs on the **Jobs** page of the console, and therefore won't be able to choose the job to display its details. For the details to include information about a custom data identifier that the job uses, the user must also be allowed to perform the `macie2:BatchGetCustomDataIdentifiers` action for the custom data identifier.

Example: Allow users to review their own permissions

This example shows how you might create a policy that allows IAM users to view the inline and managed policies that are attached to their user identity. This policy includes permissions to complete this action on the console or programmatically using the AWS CLI or AWS API.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "ViewOwnUserInfo",
            "Effect": "Allow",
            "Action": [
                "iam:GetUserPolicy",
                "iam:ListGroupsForUser",
                "iam:ListAttachedUserPolicies",
                "iam:ListUserPolicies",
                "iam:GetUser"
            ],
            "Resource": ["arn:aws:iam::*:user/${aws:username}"]
        },
        {
            "Sid": "NavigateInConsole",
            "Effect": "Allow",
            "Action": [
                "iam:GetGroupPolicy",
                "iam:GetPolicyVersion",
                "iam:GetPolicy",
                "iam:ListAttachedGroupPolicies",
                "iam:ListGroupPolicies",
                "iam:ListPolicyVersions",
                "iam:ListPolicies",
                "iam:ListUsers"
            ],
            "Resource": "*"
```

```
        }
    ]
}
```

Example: Allow users to create sensitive data discovery jobs

This example shows how you might create a policy that allows a user to create sensitive data discovery jobs.

In the example, the first statement grants `macie2:CreateClassificationJob` permissions to the user. These permissions allow the user to create jobs. The statement also grants `macie2:DescribeClassificationJob` permissions. These permissions allow the user to access the details of existing jobs. Although these permissions aren't required to create jobs, access to these details can help the user create jobs that have unique configuration settings.

The second statement in the example allows the user to create, configure, and review jobs by using the Amazon Macie console. The `macie2:ListClassificationJobs` permissions allow the user to display existing jobs on the **Jobs** page of the console. All other permissions in the statement allow the user to configure and create a job by using the **Create job** pages on the console.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "CreateAndReviewJobs",
            "Effect": "Allow",
            "Action": [
                "macie2:CreateClassificationJob",
                "macie2:DescribeClassificationJob"
            ],
            "Resource": "arn:aws:macie2:*:*:classification-job/*"
        },
        {
            "Sid": "CreateAndReviewJobsOnConsole",
            "Effect": "Allow",
            "Action": [
                "macie2:ListClassificationJobs",
                "macie2:ListAllowLists",
                "macie2:ListCustomDataIdentifiers",
```

```
                "macie2:ListManagedDataIdentifiers",
                "macie2:SearchResources",
                "macie2:DescribeBuckets"
            ],
            "Resource": "*"
        }
    ]
}
```

Example: Allow users to manage a sensitive data discovery job

This example shows how you might create a policy that allows a user to access the details of a particular sensitive data discovery job, the job whose ID is `3ce05dbb7ec5505def334104bexample`. The example also allows the user to change the status of the job as necessary.

The first statement in the example grants `macie2:DescribeClassificationJob` and `macie2:UpdateClassificationJob` permissions to the user. These permissions allow the user to retrieve the job's details and change the job's status, respectively. The second statement grants `macie2:ListClassificationJobs` permissions to the user, which allows the user to access the job by using the **Jobs** page on the Amazon Macie console.

JSON

```
{
   "Version": "2012-10-17",
   "Statement": [
 {
        "Sid": "ManageOneJob",
        "Effect": "Allow",
        "Action": [
           "macie2:DescribeClassificationJob",
           "macie2:UpdateClassificationJob"
        ],
        "Resource": "arn:aws:macie2:*:*:classification-
job/3ce05dbb7ec5505def334104bexample"
 },
 {
        "Sid": "ListJobsOnConsole",
        "Effect": "Allow",
```

```
            "Action": "macie2:ListClassificationJobs",
            "Resource": "*"
        }
    ]
}
```

You might also allow the user to access logging data (*log events*) that Macie publishes to Amazon CloudWatch Logs for the job. To do this, you can add statements that grant permissions to perform CloudWatch Logs (`logs`) actions on the log group and stream for the job. For example:

```
{
    "Sid": "AccessLogGroupForMacieJobs",
    "Effect": "Allow",
    "Action": [
        "logs:DescribeLogGroups",
        "logs:DescribeLogStreams"
    ],
    "Resource": "arn:aws:logs:*:*:log-group:aws/macie/classificationjobs"
},
{
    "Sid": "AccessLogEventsForOneMacieJob",
    "Effect": "Allow",
    "Action": "logs:GetLogEvents",
    "Resource": [
        "arn:aws:logs:*:*:log-group:aws/macie/classificationjobs/*",
        "arn:aws:logs:*:*:log-group:aws/macie/classificationjobs:log-
stream:3ce05dbb7ec5505def334104bexample"
    ]
}
```

For information about managing access to CloudWatch Logs, see Overview of managing access permissions to your CloudWatch Logs resources in the *Amazon CloudWatch Logs User Guide*.

Example: Allow users to review findings

This example shows how you might create a policy that allows a user to access findings data.

In this example, the `macie2:GetFindings` and `macie2:GetFindingStatistics` permissions allow the user to retrieve the data by using the Amazon Macie API or the Amazon Macie console. The `macie2:ListFindings` permissions allow the user to retrieve and review the data by using the **Summary** dashboard and the **Findings** pages on the Amazon Macie console.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "ReviewFindings",
            "Effect": "Allow",
            "Action": [
                "macie2:GetFindings",
                "macie2:GetFindingStatistics",
                "macie2:ListFindings"
            ],
            "Resource": "*"
        }
    ]
}
```

You might also allow the user to create and manage filter rules and suppression rules for findings. To do this, you might include a statement that grants the following permissions: `macie2:CreateFindingsFilter`, `macie2:GetFindingsFilter`, `macie2:UpdateFindingsFilter`, and `macie2:DeleteFindingsFilter`. To allow the user to manage the rules by using the Amazon Macie console, also include `macie2:ListFindingsFilters` permissions in the policy. For example:

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "ReviewFindings",
            "Effect": "Allow",
            "Action": [
                "macie2:GetFindings",
                "macie2:GetFindingStatistics",
                "macie2:ListFindings"
            ],
            "Resource": "*"
        },
```

```
        {
            "Sid": "ManageRules",
            "Effect": "Allow",
            "Action": [
                "macie2:GetFindingsFilter",
                "macie2:UpdateFindingsFilter",
                "macie2:CreateFindingsFilter",
                "macie2:DeleteFindingsFilter"
            ],
            "Resource": "arn:aws:macie2:*:*:findings-filter/*"
        },
        {
            "Sid": "ListRulesOnConsole",
            "Effect": "Allow",
            "Action": "macie2:ListFindingsFilters",
            "Resource": "*"
        }
    ]
}
```

Example: Allow users to review custom data identifiers based on tags

In identity-based policies, you can use conditions to control access to Amazon Macie resources based on tags. This example shows how you might create a policy that allows a user to review custom data identifiers by using the Amazon Macie console or the Amazon Macie API. However, permission is granted only if the value for the `Owner` tag is the user's username.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "ReviewCustomDataIdentifiersIfOwner",
            "Effect": "Allow",
            "Action": "macie2:GetCustomDataIdentifier",
            "Resource": "arn:aws:macie2:*:*:custom-data-identifier/*",
            "Condition": {
                "StringEquals": {"aws:ResourceTag/Owner": "${aws:username}"}
            }
        },
```

```
        {
            "Sid": "ListCustomDataIdentifiersOnConsoleIfOwner",
            "Effect": "Allow",
            "Action": "macie2:ListCustomDataIdentifiers",
            "Resource": "*",
            "Condition": {
                "StringEquals": {"aws:ResourceTag/Owner": "${aws:username}"}
            }
        }
    ]
}
```

In this example, if a user who has the username `richard-roe` attempts to review the details of a custom data identifier, the custom data identifier must be tagged `Owner=richard-roe` or `owner=richard-roe`. Otherwise, the user is denied access. The condition tag key `Owner` matches both `Owner` and `owner` because condition key names aren't case sensitive. For more information, see IAM JSON policy elements: Condition in the *IAM User Guide*.

AWS managed policies for Macie

An AWS managed policy is a standalone policy that is created and administered by AWS. AWS managed policies are designed to provide permissions for many common use cases so that you can start assigning permissions to users, groups, and roles.

Keep in mind that AWS managed policies might not grant least-privilege permissions for your specific use cases because they're available for all AWS customers to use. We recommend that you reduce permissions further by defining customer managed policies that are specific to your use cases.

You cannot change the permissions defined in AWS managed policies. If AWS updates the permissions defined in an AWS managed policy, the update affects all principal identities (users, groups, and roles) that the policy is attached to. AWS is most likely to update an AWS managed policy when a new AWS service is launched or new API operations become available for existing services.

For more information, see AWS managed policies in the *IAM User Guide*.

Amazon Macie provides several AWS managed policies: the `AmazonMacieFullAccess` policy, the `AmazonMacieReadOnlyAccess` policy, and the `AmazonMacieServiceRolePolicy` policy.

Policies and updates

- AWS managed policy: AmazonMacieFullAccess
- AWS managed policy: AmazonMacieReadOnlyAccess
- AWS managed policy: AmazonMacieServiceRolePolicy
- Updates to AWS managed policies for Macie

AWS managed policy: AmazonMacieFullAccess

You can attach the `AmazonMacieFullAccess` policy to your IAM entities.

This policy grants full administrative permissions that allow an IAM identity (*principal*) to create the Amazon Macie service-linked role and perform all read and write actions for Amazon Macie. The permissions include mutating functions such as create, update, and delete. If this policy is attached to a principal, the principal can create, retrieve, and otherwise access all Macie resources, data, and settings for their account.

This policy must be attached to a principal before the principal can enable Macie for their account —a principal must be allowed to create the Macie service-linked role in order to enable Macie for their account.

Permissions details

This policy includes the following permissions:

- `macie2` – Allows principals to perform all read and write actions for Amazon Macie.
- `iam` – Allows principals to create service-linked roles. The `Resource` element specifies the service-linked role for Macie. The `Condition` element uses the `iam:AWSServiceName` condition key and the `StringLike` condition operator to restrict permissions to the service-linked role for Macie.
- `pricing` – Allows principals to retrieve pricing data for their AWS account from AWS Billing and Cost Management. Macie uses this data to calculate and display estimated costs when principals create and configure sensitive data discovery jobs.

To review the permissions for this policy, see AmazonMacieFullAccess in the *AWS Managed Policy Reference Guide*.

AWS managed policy: AmazonMacieReadOnlyAccess

You can attach the `AmazonMacieReadOnlyAccess` policy to your IAM entities.

This policy grants read-only permissions that allow an IAM identity (*principal*) to perform all read actions for Amazon Macie. The permissions don't include mutating functions such as create, update, or delete. If this policy is attached to a principal, the principal can retrieve but not otherwise access all Macie resources, data, and settings for their account.

Permissions details

This policy includes the following permissions:

`macie2` – Allows principals to perform all read actions for Amazon Macie.

To review the permissions for this policy, see AmazonMacieReadOnlyAccess in the *AWS Managed Policy Reference Guide*.

AWS managed policy: AmazonMacieServiceRolePolicy

You can't attach the `AmazonMacieServiceRolePolicy` policy to your IAM entities.

This policy is attached to a service-linked role that allows Amazon Macie to perform actions on your behalf. For more information, see Using service-linked roles for Macie.

To review the permissions for this policy, see AmazonMacieServiceRolePolicy in the *AWS Managed Policy Reference Guide*.

Updates to AWS managed policies for Macie

The following table provides details about updates to AWS managed policies for Amazon Macie since this service began tracking these changes. For automatic alerts about updates to the policies, subscribe to the RSS feed on the Macie document history page.

Change	Description	Date
AmazonMacieReadOnlyAccess – Added a new policy	Macie added a new policy, the `AmazonMacieReadOnlyAccess` policy. This policy grants read-only permissions that allow principals to retrieve all Macie resources, data, and settings for their account.	June 15, 2023
AmazonMacieFullAccess – Updated an existing policy	In the `AmazonMacieFullAccess` policy, Macie updated the Amazon Resource Name (ARN) of the Macie service-linked role (`aws-service-role/macie.amazonaws.com/AWSServiceRoleForAmazonMacie`).	June 30, 2022
AmazonMacieServiceRolePolicy – Updated an existing policy	Macie removed actions and resources for Amazon Macie Classic from the `AmazonMacieServiceRolePolicy` policy. Amazon Macie Classic has been discontinued and is no longer available. More specifically, Macie removed all AWS CloudTrail actions. Macie also removed all Amazon S3 actions for	May 20, 2022

Change	Description	Date
	the following resources: `arn:aws:s3:::awsmacie-*` , `arn:aws:s3:::awsmacietrail-*` , and `arn:aws:s3:::*-awsmacietrail-*` .	
AmazonMacieFullAccess – Updated an existing policy	Macie added an AWS Billing and Cost Management (`pricing`) action to the `AmazonMacieFullAccess` policy. This action allows principals to retrieve pricing data for their account. Macie uses this data to calculate and display estimated costs when principals create and configure sensitive data discovery jobs. Macie also removed Amazon Macie Classic (`macie`) actions from the `AmazonMacieFullAccess` policy.	March 7, 2022
AmazonMacieServiceRolePolicy – Updated an existing policy	Macie added Amazon CloudWatch Logs actions to the `AmazonMacieServiceRolePolicy` policy. These actions allow Macie to publish log events to CloudWatch Logs for sensitive data discovery jobs.	April 13, 2021

Change	Description	Date
Macie started tracking changes	Macie started tracking changes for its AWS managed policies.	April 13, 2021

Using service-linked roles for Macie

Amazon Macie uses an AWS Identity and Access Management (IAM) service-linked role named `AWSServiceRoleForAmazonMacie`. This service-linked role is an IAM role that's linked directly to Macie. It's predefined by Macie and it includes all the permissions that Macie requires to call other AWS services and monitor AWS resources on your behalf. Macie uses this service-linked role in all the AWS Regions where Macie is available.

A service-linked role makes setting up Macie easier because you don't have to manually add the necessary permissions. Macie defines the permissions of this service-linked role, and unless defined otherwise, only Macie can assume the role. The defined permissions include the trust policy and the permissions policy, and that permissions policy can't be attached to any other IAM entity.

For information about other services that support service-linked roles, see AWS services that work with IAM and look for the services that have **Yes** in the **Service-linked roles** column. Choose a **Yes** with a link to review the service-linked role documentation for that service.

Topics

- Service-linked role permissions for Macie
- Creating the service-linked role for Macie
- Editing the service-linked role for Macie
- Deleting the service-linked role for Macie
- Supported AWS Regions for the Macie service-linked role

Service-linked role permissions for Macie

Amazon Macie uses the service-linked role named `AWSServiceRoleForAmazonMacie`. This service-linked role trusts the `macie.amazonaws.com` service to assume the role.

The permissions policy for the role, which is named `AmazonMacieServiceRolePolicy`, allows Macie to perform tasks such as the following on the specified resources:

- Use Amazon S3 actions to retrieve information about S3 buckets and objects.
- Use Amazon S3 actions to retrieve S3 objects.
- Use AWS Organizations actions to retrieve information about associated accounts.
- Use Amazon CloudWatch Logs actions to log events for sensitive data discovery jobs.

To review the permissions for this policy, see AmazonMacieServiceRolePolicy in the *AWS Managed Policy Reference Guide*.

For details about updates to this policy, see Updates to AWS managed policies for Macie. For automatic alerts about changes to this policy, subscribe to the RSS feed on the Macie document history page.

You must configure permissions for an IAM entity (such as a user or role) to allow the entity to create, edit, or delete a service-linked role. For more information, see Service-linked role permissions in the *IAM User Guide*.

Creating the service-linked role for Macie

You don't need to manually create the `AWSServiceRoleForAmazonMacie` service-linked role for Amazon Macie. When you enable Macie for your AWS account, Macie automatically creates the service-linked role for you.

If you delete the Macie service-linked role and then need to create it again, you can use the same process to re-create the role in your account. When you enable Macie again, Macie creates the service-linked role again for you.

Editing the service-linked role for Macie

Amazon Macie doesn't allow you to edit the `AWSServiceRoleForAmazonMacie` service-linked role. After a service-linked role is created, you can't change the name of the role because various entities might reference the role. However, you can edit the description of the role by using IAM. For more information, see Updating a service-linked role in the *IAM User Guide*.

Deleting the service-linked role for Macie

You can delete a service-linked role only after you delete its related resources. This protects your resources because you can't inadvertently remove permission to access the resources.

If you no longer need to use Amazon Macie, we recommend that you manually delete the `AWSServiceRoleForAmazonMacie` service-linked role. When you disable Macie, Macie doesn't delete the role for you.

Before you delete the role, you must disable Macie in each AWS Region where you enabled it. You must also manually clean up the resources for the role. To delete the role, you can use the IAM console, the AWS CLI, or the AWS API. For more information, see Deleting a service-linked role in the *IAM User Guide*.

> **Note**
>
> If Macie is using the `AWSServiceRoleForAmazonMacie` role when you try to delete the resources, the deletion might fail. If that happens, wait a few minutes and then try the operation again.

If you delete the `AWSServiceRoleForAmazonMacie` service-linked role and need to create it again, you can create it again by enabling Macie for your account. When you enable Macie again, Macie creates the service-linked role again for you.

Supported AWS Regions for the Macie service-linked role

Amazon Macie supports using the `AWSServiceRoleForAmazonMacie` service-linked role in all the AWS Regions where Macie is available. For a list of Regions where Macie is currently available, see Amazon Macie endpoints and quotas in the *AWS General Reference*.

Troubleshooting identity and access management for Macie

The following information can help you diagnose and fix common issues that you might encounter when working with Amazon Macie and AWS Identity and Access Management (IAM).

Topics

- I'm not authorized to perform an action in Macie
- I want to allow people outside my AWS account to access my Macie resources

I'm not authorized to perform an action in Macie

If you receive an error that you're not authorized to perform an action, your policies must be updated to allow you to perform the action.

The following example error occurs when the `mateojackson` IAM user tries to use the console to view details about a fictional *`my-example-widget`* resource but doesn't have the fictional `macie2:`*`GetWidget`* permissions.

```
User: arn:aws:iam::123456789012:user/mateojackson is not authorized to perform:
 macie2:GetWidget on resource: my-example-widget
```

In this case, the policy for the `mateojackson` user must be updated to allow access to the *`my-example-widget`* resource by using the `macie2:`*`GetWidget`* action.

If you need help, contact your AWS administrator. Your administrator is the person who provided you with your sign-in credentials.

I want to allow people outside my AWS account to access my Macie resources

You can create a role that users in other accounts or people outside of your organization can use to access your resources. You can specify who is trusted to assume the role. For services that support resource-based policies or access control lists (ACLs), you can use those policies to grant people access to your resources.

To learn more, consult the following:

- To learn whether Macie supports these features, see How Macie works with AWS Identity and Access Management.
- To learn how to provide access to your resources across AWS accounts that you own, see Providing access to an IAM user in another AWS account that you own in the *IAM User Guide*.
- To learn how to provide access to your resources to third-party AWS accounts, see Providing access to AWS accounts owned by third parties in the *IAM User Guide*.
- To learn how to provide access through identity federation, see Providing access to externally authenticated users (identity federation) in the *IAM User Guide*.
- To learn the difference between using roles and resource-based policies for cross-account access, see Cross account resource access in IAM in the *IAM User Guide*.

Compliance validation for Macie

To learn whether an AWS service is within the scope of specific compliance programs, see AWS services in Scope by Compliance Program and choose the compliance program that you are interested in. For general information, see AWS Compliance Programs.

You can download third-party audit reports using AWS Artifact. For more information, see Downloading Reports in AWS Artifact.

Your compliance responsibility when using AWS services is determined by the sensitivity of your data, your company's compliance objectives, and applicable laws and regulations. AWS provides the following resources to help with compliance:

- Security Compliance & Governance – These solution implementation guides discuss architectural considerations and provide steps for deploying security and compliance features.
- HIPAA Eligible Services Reference – Lists HIPAA eligible services. Not all AWS services are HIPAA eligible.
- AWS Compliance Resources – This collection of workbooks and guides might apply to your industry and location.
- AWS Customer Compliance Guides – Understand the shared responsibility model through the lens of compliance. The guides summarize the best practices for securing AWS services and map the guidance to security controls across multiple frameworks (including National Institute of Standards and Technology (NIST), Payment Card Industry Security Standards Council (PCI), and International Organization for Standardization (ISO)).
- Evaluating Resources with Rules in the *AWS Config Developer Guide* – The AWS Config service assesses how well your resource configurations comply with internal practices, industry guidelines, and regulations.
- AWS Security Hub – This AWS service provides a comprehensive view of your security state within AWS. Security Hub uses security controls to evaluate your AWS resources and to check your compliance against security industry standards and best practices. For a list of supported services and controls, see Security Hub controls reference.
- Amazon GuardDuty – This AWS service detects potential threats to your AWS accounts, workloads, containers, and data by monitoring your environment for suspicious and malicious activities. GuardDuty can help you address various compliance requirements, like PCI DSS, by meeting intrusion detection requirements mandated by certain compliance frameworks.
- AWS Audit Manager – This AWS service helps you continuously audit your AWS usage to simplify how you manage risk and compliance with regulations and industry standards.

Resilience in Macie

The AWS global infrastructure is built around AWS Regions and Availability Zones. Regions provide multiple physically separated and isolated Availability Zones, which are connected through

low-latency, high-throughput, and highly redundant networking. With Availability Zones, you can design and operate applications and databases that automatically fail over between zones without interruption. Availability Zones are more highly available, fault tolerant, and scalable than traditional single or multiple data center infrastructures. For more information about AWS Regions and Availability Zones, see AWS Global Infrastructure.

In addition to the AWS global infrastructure, Amazon Macie offers several features to help support your data resiliency and backup needs. For example, when you run a sensitive data discovery job or Macie performs automated sensitive data discovery, Macie automatically creates an analysis record for each Amazon Simple Storage Service (Amazon S3) object that's included in the scope of the analysis. These records, referred to as a *sensitive data discovery results*, log details about the analysis that Macie performs on individual S3 objects. This includes objects that Macie doesn't detect sensitive data in, and objects that Macie can't analyze due to errors or issues. Macie stores these results in an S3 bucket that you specify. For more information, see Storing and retaining sensitive data discovery results.

Macie also publishes policy and sensitive data findings to Amazon EventBridge as events. This includes new findings and updates to existing policy findings. (It doesn't include findings that you archive automatically using suppression rules.) By using EventBridge, you can send findings data to your preferred storage platform and store the data for as long as you like. Depending on publication settings that you choose, Macie can also publish policy and sensitive data findings to AWS Security Hub. For more information, see *Monitoring and processing findings*.

You also have the option of using Macie API operations to retrieve findings and other types of data programmatically. You can then process and send the data to your preferred storage platform, or another service, application, or system. For information about API operations that you might use to do this, see the Amazon Macie API Reference.

Infrastructure security in Macie

As a managed service, Amazon Macie is protected by AWS global network security. For information about AWS security services and how AWS protects infrastructure, see AWS Cloud Security. To design your AWS environment using the best practices for infrastructure security, see Infrastructure Protection in *Security Pillar AWS Well-Architected Framework*.

You use AWS published API calls to access Macie through the network. Clients must support the following:

- Transport Layer Security (TLS). We require TLS 1.2 and recommend TLS 1.3.

- Cipher suites with perfect forward secrecy (PFS) such as DHE (Ephemeral Diffie-Hellman) or ECDHE (Elliptic Curve Ephemeral Diffie-Hellman). Most modern systems such as Java 7 and later support these modes.

Additionally, requests must be signed by using an access key ID and a secret access key that is associated with an IAM principal. Or you can use the AWS Security Token Service (AWS STS) to generate temporary security credentials to sign requests.

You can call these API operations from any network location. However, if you use Amazon Virtual Private Cloud (Amazon VPC) to host your AWS resources, you can establish a private connection between your VPC and Macie by creating an interface endpoint. Interface endpoints are powered by AWS PrivateLink, a technology that enables you to privately access Macie without an internet gateway, NAT device, VPN connection, or AWS Direct Connect connection. We create an endpoint network interface in each subnet that you enable for an interface endpoint. These are requester-managed network interfaces that can serve as the entry point for traffic destined for Macie. For more information, see Access AWS services through AWS PrivateLink in the *AWS PrivateLink Guide*.

Accessing Macie with an interface endpoint (AWS PrivateLink)

You can use AWS PrivateLink to create a private connection between your virtual private cloud (VPC) and Amazon Macie. You can access Macie as if it were in your VPC, without the use of an internet gateway, NAT device, VPN connection, or AWS Direct Connect connection. Instances in your VPC don't need public IP addresses to access Macie.

You establish this private connection by creating an *interface endpoint*, powered by AWS PrivateLink. We create an endpoint network interface in each subnet that you enable for the interface endpoint. These are requester-managed network interfaces that serve as the entry point for traffic destined for Macie.

For more information, see Access AWS services through AWS PrivateLink in the *AWS PrivateLink Guide*.

Topics

- Considerations for Macie interface endpoints
- Creating an interface endpoint for Macie
- Creating an endpoint policy for Macie

Considerations for Macie interface endpoints

Amazon Macie supports interface endpoints in all the AWS Regions where it's currently available. For a list of these Regions, see Amazon Macie endpoints and quotas in the *AWS General Reference*. Macie supports making calls to all of its API operations through interface endpoints.

If you create an interface endpoint for Macie, consider doing the same for other AWS services that integrate with Macie and with AWS PrivateLink, such as Amazon EventBridge and AWS Security Hub. Macie and those services can then use the interface endpoints for the integration. For example, if you create an interface endpoint for Macie and an interface endpoint for Security Hub, Macie can use its interface endpoint when it publishes findings to Security Hub. Security Hub can use its interface endpoint when it receives the findings. For information about supported services, see AWS services that integrate with AWS PrivateLink in the *AWS PrivateLink Guide*.

Creating an interface endpoint for Macie

You can create an interface endpoint for Amazon Macie by using the Amazon VPC console or the AWS Command Line Interface (AWS CLI). For more information, see Create a VPC endpoint in the *AWS PrivateLink Guide*.

When you create an interface endpoint for Macie, use the following service name:

`com.amazonaws.region.macie2`

Where *region* is the Region code for the applicable AWS Region.

If you enable private DNS for the interface endpoint, you can make API requests to Macie using its default Regional DNS name, for example, `macie2.us-east-1.amazonaws.com` for the US East (N. Virginia) Region.

Creating an endpoint policy for Macie

An *endpoint policy* is an AWS Identity and Access Management (IAM) resource that you can attach to an interface endpoint. The default endpoint policy allows full access to Amazon Macie through the interface endpoint. To control the access allowed to Macie from your VPC, attach a custom endpoint policy to the interface endpoint.

An endpoint policy specifies the following information:

- The principals that can perform actions (AWS accounts, IAM users, and IAM roles).
- The actions that can be performed.
- The resources on which the actions can be performed.

It's a separate policy for controlling access from the endpoint to the specified service. For more information, see Control access to VPC endpoints using endpoint policies in the *AWS PrivateLink Guide*.

Example: VPC endpoint policy for Macie actions

The following is an example of a custom endpoint policy for Macie. If you attach this policy to your interface endpoint, it grants access to the listed Macie actions for all principals on all resources. It allows users connecting to Macie through the VPC to access findings data by using the Amazon Macie API.

```
{
   "Statement": [
      {
         "Principal": "*",
         "Effect": "Allow",
         "Action": [
            "macie2:GetFindings",
            "macie2:GetFindingStatistics",
            "macie2:ListFindings"
         ],
         "Resource": "*"
      }
   ]
}
```

To also allow users to access findings data or perform other actions by using the Amazon Macie console, the policy should also grant access to the `macie2:GetMacieSession` action, for example:

```
{
   "Statement": [
      {
         "Principal": "*",
         "Effect": "Allow",
         "Action": [
```

```
            "macie2:GetMacieSession",
            "macie2:GetFindings",
            "macie2:GetFindingStatistics",
            "macie2:ListFindings"
         ],
         "Resource": "*"
      }
   ]
}
```

Logging Macie API calls with AWS CloudTrail

Amazon Macie integrates with AWS CloudTrail, which is a service that provides a record of actions taken by a user, a role, or an AWS service. CloudTrail captures all API calls for Macie as management events. The calls captured include calls from the Amazon Macie console and programmatic calls to Amazon Macie API operations. By using the information collected by CloudTrail, you can determine the request that was made to Macie, the IP address from which the request was made, when it was made, and additional details.

Every event or log entry contains information about who generated the request. The identity information helps you determine the following:

- Whether the request was made with root user or user credentials.
- Whether the request was made on behalf of an AWS IAM Identity Center user.
- Whether the request was made with temporary security credentials for a role or federated user.
- Whether the request was made by another AWS service.

CloudTrail is active in your AWS account when you create the account and you automatically have access to the CloudTrail **Event history**. The CloudTrail **Event history** provides a viewable, searchable, downloadable, and immutable record of the past 90 days of recorded management events in an AWS Region. For more information, see Working with CloudTrail Event history in the *AWS CloudTrail User Guide*. There are no CloudTrail charges for viewing the **Event history**.

For an ongoing record of events in your AWS account past 90 days, create a trail or a CloudTrail Lake event data store.

CloudTrail trails

A *trail* enables CloudTrail to deliver log files to an Amazon S3 bucket. All trails created using the AWS Management Console are multi-Region. You can create a single-Region or a multi-Region trail by using the AWS CLI. Creating a multi-Region trail is recommended because you capture activity in all AWS Regions in your account. If you create a single-Region trail, you can view only the events logged in the trail's AWS Region. For more information about trails, see Creating a trail for your AWS account and Creating a trail for an organization in the *AWS CloudTrail User Guide*.

You can deliver one copy of your ongoing management events to your Amazon S3 bucket at no charge from CloudTrail by creating a trail, however, there are Amazon S3 storage charges. For more information about CloudTrail pricing, see AWS CloudTrail Pricing. For information about Amazon S3 pricing, see Amazon S3 Pricing.

CloudTrail Lake event data stores

CloudTrail Lake lets you run SQL-based queries on your events. CloudTrail Lake converts existing events in row-based JSON format to Apache ORC format. ORC is a columnar storage format that is optimized for fast retrieval of data. Events are aggregated into *event data stores*, which are immutable collections of events based on criteria that you select by applying advanced event selectors. The selectors that you apply to an event data store control which events persist and are available for you to query. For more information about CloudTrail Lake, see Working with AWS CloudTrail Lake in the *AWS CloudTrail User Guide*.

CloudTrail Lake event data stores and queries incur costs. When you create an event data store, you choose the pricing option you want to use for the event data store. The pricing option determines the cost for ingesting and storing events, and the default and maximum retention period for the event data store. For more information about CloudTrail pricing, see AWS CloudTrail Pricing.

Macie management events in AWS CloudTrail

Management events provide information about management operations that are performed on resources in your AWS account. These are also known as control plane operations. By default, CloudTrail logs management events.

Amazon Macie logs all Macie control plane operations as management events in CloudTrail. For example, calls to the `ListFindings`, `DescribeBuckets`, and `CreateClassificationJob` operations generate management events in CloudTrail. Each event includes an `eventSource` field. This field indicates the AWS service that a request was made to. For Macie events, the value for this field is: `macie2.amazonaws.com`.

For a list of the control plane operations that Macie logs in CloudTrail, see Operations in the *Amazon Macie API Reference*.

Examples of Macie events in AWS CloudTrail

An event represents a single request from any source and includes information about the requested API operation, the date and time of the operation, request parameters, and so on. CloudTrail log files aren't an ordered stack trace of the public API calls, so events don't appear in any specific order.

The following examples show CloudTrail events that demonstrate Amazon Macie operations. For details about the information that an event might contain, see CloudTrail record contents in the *AWS CloudTrail User Guide*.

Example: Listing findings

The following example shows a CloudTrail event for the Amazon Macie ListFindings operation. In this example, an AWS Identity and Access Management (IAM) user (`Mary_Major`) used the Amazon Macie console to retrieve a subset of information about current policy findings for their account.


```
{
    "eventVersion": "1.08",
    "userIdentity": {
        "type": "IAMUser",
        "principalId": "123456789012",
        "arn": "arn:aws:iam::123456789012:user/Mary_Major",
        "accountId": "123456789012",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE",
        "userName": "Mary_Major",
        "sessionContext":{
            "attributes": {
                "creationdate": "2023-11-14T15:49:57Z",
                "mfaAuthenticated": "false"
            }
        }
    },
    "eventTime": "2023-11-14T16:09:56Z",
    "eventSource": "macie2.amazonaws.com",
    "eventName": "ListFindings",
    "awsRegion": "us-east-1",
    "sourceIPAddress": "198.51.100.1",
    "userAgent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML,
 like Gecko) Chrome/119.0.0.0 Safari/537.36",
    "requestParameters": {
        "sortCriteria": {
```

```
            "attributeName": "updatedAt",
            "orderBy": "DESC"
        },
        "findingCriteria": {
            "criterion": {
                "archived": {
                    "eq": [
                        "false"
                    ]
                },
                "category": {
                    "eq": [
                        "POLICY"
                    ]
                }
            }
        },
        "maxResults": 25,
        "nextToken": ""
    },
    "responseElements": null,
    "requestID": "d58af6be-1115-4a41-91f8-ace03example",
    "eventID": "ad97fac5-f7cf-4ff9-9cf2-d0676example",
    "readOnly": true,
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "123456789012",
    "eventCategory": "Management"
}
```


Example: Retrieving sensitive data samples for a finding

This example shows CloudTrail events for retrieving and revealing samples of sensitive data that Amazon Macie reported in a finding. In this example, an AWS Identity and Access Management (IAM) user (`JohnDoe`) used the Amazon Macie console to retrieve and reveal sensitive data samples. The user's account is configured to assume an IAM role (`MacieReveal`) to retrieve and reveal sensitive data samples from affected Amazon Simple Storage Service (Amazon S3) objects.

The following event shows details about the user's request to retrieve and reveal sensitive data samples by performing the Amazon Macie GetSensitiveDataOccurrences operation.


```
{
```

```
    "eventVersion": "1.08",
    "userIdentity": {
        "type": "AssumedRole",
        "principalId": "UU4MH7OYK5ZCOAEXAMPLE:JohnDoe",
        "arn": "arn:aws:sts::111122223333:assumed-role/Admin/JohnDoe",
        "accountId": "111122223333",
        "accessKeyId": "AKIAI44QH8DHBEXAMPLE",
        "sessionContext": {
            "sessionIssuer": {
                "type": "Role",
                "principalId": "UU4MH7OYK5ZCOAEXAMPLE",
                "arn": "arn:aws:iam::111122223333:role/Admin",
                "accountId": "111122223333",
                "userName": "Admin"
            },
            "webIdFederationData": {},
            "attributes": {
                "creationDate": "2023-12-12T14:40:23Z",
                "mfaAuthenticated": "false"
            }
        }
    },
    "eventTime": "2023-12-12T17:04:47Z",
    "eventSource": "macie2.amazonaws.com",
    "eventName": "GetSensitiveDataOccurrences",
    "awsRegion": "us-east-1",
    "sourceIPAddress": "198.51.100.252",
    "userAgent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML,
 like Gecko) Chrome/119.0.0.0 Safari/537.36",
    "requestParameters": {
        "findingId": "3ad9d8cd61c5c390bede45cd2example"
    },
    "responseElements": null,
    "requestID": "c30cb760-5102-47e7-88d8-ff2e8example",
    "eventID": "baf52d92-f9c3-431a-bfe8-71c81example",
    "readOnly": true,
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "111122223333",
    "eventCategory": "Management"
}
```

The next event shows details about Macie then assuming the specified IAM role (`MacieReveal`) by performing the AWS Security Token Service (AWS STS) AssumeRole operation.


```
{
    "eventVersion": "1.08",
    "userIdentity": {
        "type": "AWSService",
        "invokedBy": "reveal-samples.macie.amazonaws.com"
    },
    "eventTime": "2023-12-12T17:04:47Z",
    "eventSource": "sts.amazonaws.com",
    "eventName": "AssumeRole",
    "awsRegion": "us-east-1",
    "sourceIPAddress": "reveal-samples.macie.amazonaws.com",
    "userAgent": "reveal-samples.macie.amazonaws.com",
    "requestParameters": {
        "roleArn": "arn:aws:iam::111122223333:role/MacieReveal",
        "roleSessionName": "RevealCrossAccount"
    },
    "responseElements": {
        "credentials": {
            "accessKeyId": "AKIAI44QH8DHBEXAMPLE",
            "sessionToken": "XXYYaz...
EXAMPLE_SESSION_TOKEN
XXyYaZAz",
            "expiration": "Dec 12, 2023, 6:04:47 PM"
        },
        "assumedRoleUser": {
            "assumedRoleId": "AROAXOTKAROCSNEXAMPLE:RevealCrossAccount",
            "arn": "arn:aws:sts::111122223333:assumed-role/MacieReveal/
RevealCrossAccount"
        }
    },
    "requestID": "d905cea8-2dcb-44c1-948e-19419example",
    "eventID": "74ee4d0c-932d-3332-87aa-8bcf3example",
    "readOnly": true,
    "resources": [
        {
            "accountId": "111122223333",
            "type": "AWS::IAM::Role",
            "ARN": "arn:aws:iam::111122223333:role/MacieReveal"
        }
    ],
```

```
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "111122223333",
    "eventCategory": "Management"
}
```

For information about the contents of CloudTrail events, see CloudTrail record contents in the *AWS CloudTrail User Guide*.

Creating Macie resources with AWS CloudFormation

Amazon Macie integrates with AWS CloudFormation, which is a service that helps you model and set up your AWS resources so that you can spend less time creating and managing your resources and infrastructure. You create a template that describes all the AWS resources that you want (such as custom data identifiers), and AWS CloudFormation provisions and configures those resources for you.

When you use AWS CloudFormation, you can reuse your template to set up your Macie resources consistently and repeatedly. Describe your resources once, and then provision the same resources over and over in multiple AWS accounts and AWS Regions.

Macie and AWS CloudFormation templates

To provision and configure resources for Amazon Macie and related services, you must understand AWS CloudFormation templates. The templates describe the resources that you want to provision in your AWS CloudFormation stacks. They are text files in JSON or YAML format. If you're unfamiliar with JSON or YAML, AWS Infrastructure Composer or AWS CloudFormation Designer can help you get started. For more information, see Working with CloudFormation templates in the *AWS CloudFormation User Guide*.

You can create AWS CloudFormation templates for the following types of Macie resources:

- Allow lists
- Custom data identifiers
- Filter rules and suppression rules for findings, also referred to as *findings filters*

For more information, including examples of JSON and YAML templates for these types of resources, see the Amazon Macie resource type reference in the *AWS CloudFormation User Guide*.

Additional learning resources for AWS CloudFormation

To learn more about AWS CloudFormation, refer to the following resources:

- AWS CloudFormation
- AWS CloudFormation User Guide

- AWS CloudFormation API Reference
- AWS CloudFormation Command Line Interface User Guide

Suspending Macie for your AWS account

You can temporarily pause Amazon Macie for your AWS account in an AWS Region. You can do this by suspending Macie in the Region. Macie then stops performing all activities for your account in that Region. The activities include: monitoring your Amazon Simple Storage Service (Amazon S3) data, performing automated sensitive data discovery, and running sensitive data discovery jobs that are currently in progress. Macie also cancels all of your sensitive data discovery jobs in the Region. You aren't charged for using Macie in the Region while it's suspended.

If you suspend Macie in a Region, Macie retains the session identifier, settings, and resources that it stores or maintains for your account in the Region. Macie also retains certain data that it stores or maintains for your account in the Region. For example, your existing findings remain intact and are retained for up to 90 days. If automated sensitive data discovery was enabled for your account, your existing results also remain intact and are retained for up to 30 days.

Note

If your account is part of an organization that centrally manages multiple Macie accounts, note the following requirements for suspending Macie:

- If you have a member account in an AWS Organizations organization, you must contact the Macie administrator for your organization. Only your Macie administrator can suspend Macie for your account.
- If you're the Macie administrator for the organization, you must remove all member accounts that are associated with your account before you suspend Macie for your account. How you do this depends on whether your account is associated with the accounts through AWS Organizations or by invitation. For more information, see Managing multiple accounts.

After you suspend Macie in a Region, you can enable it again later. You then regain access to your Macie settings, resources, and data in the Region. In addition, Macie resumes its activities for your account in the Region. This includes updating and maintaining information about your S3 buckets, and monitoring the buckets for security and access control. This doesn't include resuming or restarting your sensitive data discovery jobs. Sensitive data discovery jobs can't be resumed or restarted after they're cancelled.

To suspend Macie for your account

To suspend Macie for your account, you can use the Amazon Macie console or the Amazon Macie API. Follow these steps to suspend it by using the console. To suspend it programmatically, use the UpdateMacieSession operation of the Amazon Macie API.

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to suspend Macie.
3. In the navigation pane, choose **Settings**.
4. In the **Suspend Macie** section, choose **Suspend Macie**.
5. When prompted for confirmation, enter **Suspend**, and then choose **Suspend**.
6. To suspend Macie in additional Regions, repeat steps 2 through 5 in each additional Region.

To subsequently re-enable Macie in a Region, open the Amazon Macie console and choose the Region by using the AWS Region selector. Then choose **Settings** in the navigation pane. In the **Suspend Macie** section, choose **Re-enable Macie**. You can also re-enable Macie programmatically. To do this, use the UpdateMacieSession operation of the Amazon Macie API.

Disabling Macie for your AWS account

If you want to stop using Amazon Macie in a particular AWS Region, you can disable it for your AWS account in the Region.

When you disable Macie in a Region, Macie stops performing all activities for your account in the Region. The activities include: monitoring your Amazon Simple Storage Service (Amazon S3) data, performing automated sensitive data discovery, and running sensitive data discovery jobs that are currently in progress. Macie also deletes all existing settings, resources, and data that it stores or maintains for your account in the Region. For example, Macie deletes your findings and sensitive data discovery jobs. Data that you stored or published to other AWS services remains intact and isn't affected—for example, sensitive data discovery results in Amazon S3 and finding events in Amazon EventBridge.

If your account is part of an organization that centrally manages multiple Macie accounts, you must do the following before you disable Macie for your account:

- If you have a member account, work with your Macie administrator to remove your account as a member account.
- If you're the Macie administrator for the organization, remove all member accounts that are associated with your account. Also delete the associations between your account and those accounts.

How you complete the preceding tasks depends on whether your account is associated with other accounts through AWS Organizations or by invitation. For more information, see Managing multiple accounts.

To disable Macie for your account

To disable Macie for your account, you can use the Amazon Macie console or the Amazon Macie API. Follow these steps to disable it by using the console. To disable it programmatically, use the DisableMacie operation of the Amazon Macie API.

> **Warning**
>
> If you disable Macie in a Region, you also permanently delete all of your existing findings, sensitive data discovery jobs, custom data identifiers, and other resources and data that Macie stores or maintains for your account in the Region. The resources and data can't be

recovered after they're deleted. To keep the resources and data, suspend Macie instead of disabling it.

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, choose the Region in which you want to disable Macie.
3. In the navigation pane, choose **Settings**.
4. In the **Disable Macie** section, choose **Disable Macie**.
5. When prompted for confirmation, enter **`Disable`**, and then choose **Disable**.

To disable Macie in additional Regions, repeat the preceding steps in each additional Region.

Quotas for Macie

Your AWS account has certain default quotas, formerly referred to as *limits*, for each AWS service. These quotas are the maximum number of service resources or operations for your account. This topic lists the quotas that apply to Amazon Macie resources and operations for your account. Unless otherwise noted, each quota applies to your account in each AWS Region.

Some quotas can be increased, while others cannot. To request an increase to a quota, use the Service Quotas console. To learn how to request an increase, see Requesting a quota increase in the *Service Quotas User Guide*. If a quota isn't available on the Service Quotas console, use the service limit increase form on the AWS Support Center Console to request an increase to the quota.

Findings

- Filter rules and suppression rules per account: 1,000
- Findings per run of a sensitive data discovery job: 100,000 + 5% of any remaining findings after the 100,000 threshold is met

  This quota applies only to the Amazon Macie console and the Amazon Macie API. There isn't a quota for the number of finding events that Macie publishes to Amazon EventBridge or the number of sensitive data discovery results that Macie creates for each run of a job.
- Detection locations per sensitive data finding: 15
- Requests to retrieve and reveal sensitive data samples from an Amazon S3 object: 100 per day

  This quota resets every 24 hours at 00:00:01 UTC+0.
- Size of an Amazon S3 object to retrieve and reveal sensitive data samples from:
  - Apache Avro object container (.avro) file: 70 MB
  - Apache Parquet (.parquet) file: 100 MB
  - CSV (.csv) file: 255 MB
  - GNU Zip compressed archive (.gz or .gzip) file: 90 MB
  - JSON or JSON Lines (.json or .jsonl) file: 25 MB
  - Microsoft Excel workbook (.xlsx) file: 20 MB
  - Non-binary text (*text/plain*) file: 100 MB
  - TSV (.tsv) file: 75 MB
  - ZIP compressed archive (.zip) file: 355 MB

If a finding applies to an archive file that generates multiple .gz files for the corresponding sensitive data discovery results, sensitive data samples can't be retrieved and revealed from the archive file.

Organizations

- Member accounts by invitation: 1,000
- Member accounts through AWS Organizations: 10,000

Preventative control monitoring

- S3 buckets per account: 10,000

  If your account exceeds this quota, Macie provides full monitoring functionality for the 10,000 buckets that were most recently created or changed. For all other buckets, Macie doesn't evaluate or monitor the buckets for security and access control, generate policy findings, or maintain complete inventory data.

Sensitive data discovery

- Monthly analysis per account by sensitive data discovery jobs: 5 TB

  This quota applies only to sensitive data discovery jobs. To increase the quota to as much as 1,000 TB (1 PB), use the Service Quotas console. To request an increase for more than 1 PB, use the service limit increase form on the AWS Support Center Console.
- Custom data identifiers per account: 10,000
- Allow lists per account: 10, 1–5 allow lists that specify predefined text and 1–5 allow lists that specify regular expressions

  Additional quotas apply to an allow list that specifies predefined text. The list can't contain more than 100,000 entries and the storage size of the list can't exceed 35 MB.
- S3 buckets to exclude from automated sensitive data discovery: 1,000

  If your account is the Macie administrator account for an organization, this quota applies to your organization overall.
- S3 buckets per sensitive data discovery job: 1,000

This quota doesn't apply to jobs that use runtime bucket criteria to determine which buckets to analyze. It applies to a job only if you configure the job to analyze specific buckets that you select. If your account is the Macie administrator account for an organization, you can select as many as 1,000 buckets spanning as many as 1,000 accounts in your organization.

- Custom data identifiers per sensitive data discovery job: 30
- Allow lists per sensitive data discovery job: 10, 1–5 allow lists that specify predefined text and 1–5 allow lists that specify regular expressions
- CreateClassificationJob operation: 0.1 requests per second
- Time to analyze an individual file: 10 hours
- Size of an individual file to analyze:
  - Adobe Portable Document Format (.pdf) file: 1,024 MB
  - Apache Avro object container (.avro) file: 8 GB
  - Apache Parquet (.parquet) file: 8 GB
  - Email message (.eml) file: 20 GB
  - GNU Zip compressed archive (.gz or .gzip) file: 8 GB
  - Microsoft Excel workbook (.xls or .xlsx) file: 512 MB
  - Microsoft Word document (.doc or .docx) file: 512 MB
  - Non-binary text file: 20 GB
  - TAR archive (.tar) file: 20 GB
  - ZIP compressed archive (.zip) file: 8 GB

  If a file is larger than the applicable quota, Macie doesn't analyze any data in the file.
- Extraction and analysis of data in a compressed or archive file:
  - Storage size (compressed): 8 GB for a GNU Zip compressed archive (.gz or .gzip) file or ZIP compressed archive (.zip) file; 20 GB for a TAR archive (.tar) file
  - Nested archive depth: 10 levels
  - Extracted files: 1,000,000
  - Extracted bytes: 10 GB of uncompressed data overall. 3 GB of uncompressed data for each extracted file that uses a supported file type or storage format.

  If the metadata for a compressed or archive file indicates that the file contains more than 10 nested levels or exceeds the applicable quota for storage size or extracted bytes, Macie

doesn't extract or analyze any data in the file. If Macie begins to extract and analyze data in a compressed or archive file and subsequently determines that the file contains more than 1,000,000 files or exceeds the quota for extracted bytes, Macie stops analyzing data in the file and creates sensitive data findings and discovery results only for the data that was processed.

- Analysis of nested elements in structured data: 256 levels per file

  This quota applies only to JSON (.json) and JSON Lines (.jsonl) files. If the nested depth of either type of file exceeds this quota, Macie doesn't analyze any data in the file.
- Detection locations per sensitive data discovery result: 1,000 per sensitive data detection type
- Detection of full names: 1,000 per file, including archive files

  After Macie detects the first 1,000 occurrences of full names in a file, Macie stops incrementing the count and reporting location data for full names.
- Detection of mailing addresses: 1,000 per file, including archive files

  After Macie detects the first 1,000 occurrences of mailing addresses in a file, Macie stops incrementing the count and reporting location data for mailing addresses.

Document history for the Amazon Macie User Guide

The following table describes the important changes to the documentation since the last release of Amazon Macie. For notification about updates to this documentation, you can subscribe to an RSS feed.

Latest documentation update: July 2, 2025

Change	Description	Date
New functionality	Macie now supports VPC interface endpoints and endpoint policies in all the AWS Regions where it's currently available.	July 2, 2025
New functionality	Macie now provides managed data identifiers that are designed to detect the following types of sensitive data: national identification numbers for Argentina, Chile, Colombia, and Mexico; Sistema Único de Boleto Electrónico (SUBE) card numbers for Argentina; and, taxpayer identification and reference numbers for Argentina, Chile, Colombia, and Mexico.	March 3, 2025
Updated functionality	Macie can now perform preventative control monitoring for up to 10,000 Amazon S3 general purpose buckets for your account.	December 6, 2024

New content	Added examples and details that explain how to configure and manage automated sensitive data discovery programmatically with the Amazon Macie API.	November 22, 2024
New feature	If you have a member account in an organization, you now have read access to statistic s, inventory data, and other information that automated sensitive data discovery produces for your Amazon S3 data. For details about the automated discovery settings for your account and organization, contact your Macie administrator.	July 22, 2024
New feature	If you're the delegated Macie administrator for an organizat ion, you can now enable or disable automated sensitive data discovery for individual accounts in your organization. With this additional option, you can now define the scope of the analyses in several ways: enable automated discovery for all accounts, selectively enable automated discovery for particular ac counts, and exclude particular S3 buckets.	June 14, 2024

New functionality	AWS Security Hub now provides security controls that check the status of Macie and automated sensitive data discovery for accounts. If these controls are enabled, Security Hub periodically runs security checks to determine whether Macie is enabled for an AWS account (Macie.1 control), and whether automated sensitive data discovery is enabled for a Macie account (Macie.2 control).	February 20, 2024
New functionality	Macie can now analyze Amazon S3 objects that are encrypted using dual-layer server-side encryption with AWS KMS keys (DSSE-KMS). These objects are now eligible for analysis when Macie performs automated sensitive data discovery or you run sensitive data discovery jobs. In addition, S3 buckets and objects that use DSSE-KMS encryption are now included in statistics and metadata that Macie provides about your Amazon S3 data.	January 17, 2024

New feature	You can now configure Macie to assume an AWS Identity and Access Management (IAM) role when you choose to retrieve and reveal samples of sensitive data that Macie reports in findings. The samples can help you verify the nature of the sensitive data that Macie found, and tailor your investigation of an affected Amazon S3 object and bucket.	November 16, 2023
New functionality	Macie now provides managed data identifiers that are designed to detect International Bank Account Numbers (IBANs) for 47 additional countries and regions. You can now use Macie to detect and report occurrences of IBANs for more than 50 countries and regions.	November 1, 2023
New functionality	Macie now provides managed data identifiers that are designed to detect the following types of sensitive data: Google Cloud API keys, Stripe API keys, and Aadhaar numbers, Permanent Account Numbers (PANs), and driver's license identification numbers for India.	September 25, 2023

New quotas	To help you verify the nature of sensitive data reported by findings, we increased the size quotas for retrieving and revealing sensitive data samples from Amazon S3 objects. You can now retrieve and reveal samples from S3 objects whose storage size exceeds 10 MB. For a list of the new quotas, see Amazon Macie quotas.	September 7, 2023
Regional availability	Macie is now available in the Israel (Tel Aviv) Region. For a complete list of AWS Regions where Macie is currently available, see Amazon Macie endpoints and quotas in the *AWS General Reference*.	August 28, 2023
Updated functionality	We implemented a new, dynamic set of default managed data identifiers for automated sensitive data discovery. The default set includes the managed data identifiers that we recommend for automated sensitive data discovery. It's designed to detect common categories and types of sensitive data while also optimizing your automated sensitive data discovery results.	August 2, 2023

Updated functionality	To help you locate occurrences of sensitive data that Macie reports in sensitive data findings and sensitive data discovery results, we changed the character limit from 20 to 240 for the names of JSON path elements in `Record` objects. This change affects new sensitive data findings and discovery results for Apache Avro object containers, Apache Parquet files, JSON files, and JSON Lines files.	July 24, 2023
Updated functionality	If you're the delegated Macie administrator for an organization in AWS Organizations, you can now manage Macie for up to 10,000 accounts in your organization.	June 30, 2023
New feature	You can now create and configure sensitive data discovery jobs to automatically use the set of managed data identifiers that we recommend for jobs. This recommended set of managed data identifiers is designed to detect common categories and types of sensitive data while also optimizing your job results.	June 28, 2023

New policy	We added a new AWS managed policy, the `AmazonMacieReadOnlyAccess` policy. This policy grants read-only permissions that allow an IAM identity (*principal*) to retrieve all Macie resources, data, and settings for their account.	June 15, 2023
New feature	To help you assess and monitor automated sensitive data discovery coverage of your Amazon S3 data, the Macie console now includes a **Resource coverage** page. The page provides a unified view of coverage statistics and data for all of your S3 buckets, including a rollup of analysis issues (if any) that recently occurred for each bucket. If issues occurred, the page also provides remediation guidance.	May 15, 2023

New feature	Macie integrates with AWS User Notifications, which is a new AWS service that acts as a central location for your AWS notifications on the AWS Management Console. With User Notifications, you can configure custom rules and delivery channels for generating and sending notifications about Amazon EventBridge events that Macie publishes for policy and sensitive data findings.	May 5, 2023
Updated content	Updated descriptions of statistics and metadata that Macie provides about default encryption settings for S3 buckets. Also updated the description of the Policy:IAMUser/S3BucketEncryptionDisabled policy finding. Amazon S3 now automatically applies server-side encryption with Amazon S3 managed keys (SSE-S3) as the base level of encryption for objects that are added to new and existing buckets. For information about this change in Amazon S3, see Setting default server-side encryption behavior for S3 buckets in the *Amazon Simple Storage Service User Guide*.	February 27, 2023

New functionality	Macie can now generate an additional type of policy finding for an S3 bucket: `Policy:IAMUser/S3BucketSharedWithCloudFront` . This type of finding indicates that a bucket's policy was changed to allow the bucket to be shared with an Amazon CloudFront origin access identity (OAI), a CloudFront origin access control (OAC), or both. In addition, buckets that are shared with CloudFront OAIs or OACs are now considered to be *shared externally* in statistics and metadata that Macie provides about your Amazon S3 data.	February 24, 2023
New functionality	Macie now supports the Amazon S3 Glacier Instant Retrieval storage class for sensitive data discovery. S3 objects that use this storage class are now eligible for analysis when Macie performs automated sensitive data discovery or you run sensitive data discovery jobs. They're also considered *classifiable objects* in statistics and metadata that Macie provides about your Amazon S3 data.	December 21, 2022

New feature	You can now configure Macie to perform automated sensitive data discovery for your account or organization. With automated sensitive data discovery, Macie continually evaluates your Amazon S3 data and uses sampling techniques to identify, select, and analyze representative objects in your S3 buckets, inspecting the objects for sensitive data. You can evaluate analyses' results in statistics, findings, and other information that Macie provides about your Amazon S3 data.	November 28, 2022
New feature	You can now create and use allow lists to specify text and text patterns that you want Macie to ignore when it inspects Amazon S3 objects for sensitive data. By using allow lists, you can define sensitive data exceptions for your particular scenarios or environment—for example, the names of public representatives for your organization, specific phone numbers, or sample data that your organization uses for testing.	August 30, 2022

New feature	To verify the nature of sensitive data that Macie finds in S3 objects, you can now configure and use Macie to retrieve samples of sensitive data reported by findings.	July 26, 2022
Updated functionality	In the AmazonMacieFullAccess policy, we updated the Amazon Resource Name (ARN) of the Macie service-linked role (`aws-service-role/macie.amazonaws.com/AWSServiceRoleForAmazonMacie`).	June 30, 2022
Updated functionality	We updated the AmazonMacieServiceRolePolicy policy, which is the policy that's attached to the Macie service-linked role (`AWSServiceRoleForAmazonMacie`). The policy no longer specifies actions and resources for Amazon Macie Classic. Amazon Macie Classic has been discontinued and is no longer available.	May 20, 2022

New functionality	Macie now includes the `OriginType` field in sensitive data findings that it publishes to AWS Security Hub. The `OriginType` field specifies how Macie found the sensitive data that produced a finding.	May 11, 2022
Updated content	Clarified how keyword and maximum match distance settings work for custom data identifiers.	April 22, 2022
New functionality	Macie now provides managed data identifiers that are designed to detect HTTP Basic Authorization headers, HTTP cookies, and JSON Web Tokens.	April 21, 2022
New content	Added descriptions and definitions of key concepts and terms for Macie.	March 16, 2022
New functionality	To calculate and display estimated costs when you create and configure sensitive data discovery jobs, Macie now retrieves pricing data for your AWS account from AWS Billing and Cost Managemen t. To support this functiona lity, we added a Billing and Cost Management action to the AmazonMacieFullAccess policy.	March 7, 2022

New functionality	Macie now includes the `Sample` field in findings that it publishes to AWS Security Hub. The `Sample` field specifies whether a finding is a sample finding.	February 24, 2022
New content	Added information about using Amazon Virtual Private Cloud to establish a private connection between your VPC and Macie.	January 19, 2022
New functionality	You can now use the Amazon Macie console to assign and manage tags for custom data identifiers, filter and suppression rules for findings, sensitive data discovery jobs, and, if you're the Macie administrator for an organization, member accounts in your organization. A *tag* is a label that you optionally define and assign to certain types of AWS resources.	January 12, 2022
New content	Added information about using AWS Identity and Access Management to manage access to Macie.	December 20, 2021

New feature	When you create a custom data identifier, you can now define severity settings for sensitive data findings that it produces. With these settings, you can specify which severity to assign to a finding based on the number of occurrences of text that match the custom data identifier's detection criteria.	November 4, 2021
New functionality	To learn about the different types of findings that Macie provides, you can generate sample findings. Sample findings use example data and placeholder values to demonstrate the kinds of information that Macie might include in each type of finding.	October 28, 2021
New functionality	Macie now includes the `OwnerAccountId` field in findings that it publishes to AWS Security Hub. This field specifies the account ID for the AWS account that owns the affected S3 bucket.	October 27, 2021

New content	Added information about centrally managing multiple Macie accounts. You can do this in two ways, by integrating Macie with AWS Organizations or by sending membership invitations from Macie.	October 13, 2021
New functionality	Your S3 bucket inventory now indicates if a bucket's permissions settings prevent Macie from retrieving information about the bucket or the bucket's objects and evaluating and monitoring the security and privacy of the bucket's data. In addition, we updated references to AWS KMS keys and customer managed keys to reflect current terminology.	October 5, 2021
New functionality	Macie now stores policy and sensitive data findings for 90 days instead of 30 days. If Macie created or updated a finding on or after August 31, 2021, you can access the finding for up to 90 days by using the Macie console or the Macie API. In certain AWS Regions, Macie began retaining findings for 90 days as early as September 27, 2021.	October 1, 2021

New feature	When you create a sensitive data discovery job, you can now specify which managed data identifiers you want the job to use when it analyzes S3 objects. With this feature, you can tailor a job's analysis to focus on certain types of sensitive data.	September 17, 2021
New functionality	Sensitive data findings now provide additional information to help you locate sensitive data in JSON and JSON Lines files.	July 6, 2021
Updated functionality	Macie now uses the `AwsS3Bucket` resource type in findings that it publishes to AWS Security Hub. (Macie previously set this value to `AWS::S3::Bucket`.) `AwsS3Bucket` is the resource type value that's used for S3 buckets in the AWS Security Finding Format (ASFF).	June 28, 2021
New feature	When you create a sensitive data discovery job, you can now define runtime criteria that determine which S3 buckets the job analyzes. With this feature, the scope of a job's analysis can dynamically adapt to changes to your bucket inventory.	May 15, 2021

New functionality	Your S3 bucket inventory and the **Summary** dashboard now provide encryption metadata and statistics indicating whether bucket policies require server-side encryption of new objects. In addition, you can now perform on-demand refreshes of object metadata for individual buckets in your bucket inventory.	April 30, 2021
New feature	You can now use Amazon CloudWatch Logs to monitor and analyze events that occur when you run sensitive data discovery jobs. To support this feature, we added CloudWatch Logs actions to the AWS managed policy for the Macie service-linked role.	April 14, 2021
Regional availability	Macie is now available in the AWS Asia Pacific (Osaka) Region.	April 5, 2021
New feature	You can now configure Macie to publish sensitive data findings to AWS Security Hub.	March 22, 2021
New content	Added information about monitoring and forecasting Macie costs and participating in the free trial.	February 26, 2021

Updated content	We replaced the term *master account* with the term *administrator account*. An administrator account is used to centrally manage multiple accounts.	February 12, 2021
New functionality	You can now refine the scope of sensitive data discovery jobs by using S3 object prefixes in custom include and exclude criteria.	February 2, 2021
Updated content	Macie now adheres to the finding type taxonomy of the AWS Security Finding Format (ASFF) when it p ublishes policy findings to AWS Security Hub.	January 28, 2021
New content	Added information about monitoring Amazon S3 data and assessing the security and privacy of that data.	January 8, 2021
Regional availability	Macie is now available in the AWS Africa (Cape Town) Region, the AWS Europe (Milan) Region, and the AWS Middle East (Bahrain) Region.	December 21, 2020

New functionality	If your account is a Macie administrator account, you can now create and run sensitive data discovery jobs that analyze data for as many as 1,000 buckets spanning as many as 1,000 accounts in your organization.	November 25, 2020
New functionality	Your S3 bucket inventory now indicates whether you've configured any one-time or periodic sensitive data discovery jobs to analyze data in a bucket. If you have, it also provides details about the job that ran most recently.	November 23, 2020
New content	Added information about filtering findings.	November 12, 2020
New functionality	Sensitive data findings now provide additional informati on to help you locate sensi tive data in Apache Avro object containers, Apache Parquet files, and Microsoft Excel workbooks.	November 9, 2020
New feature	You can now use sensitive data findings to locate individual occurrences of sensitive data in S3 objects.	October 22, 2020
New feature	You can now pause and resume sensitive data discovery jobs.	October 16, 2020

New content	Added details about the severity scoring system for policy findings and sensitive data findings.	October 6, 2020
New features	You can now view statistics that indicate how much data Macie can analyze in individual S3 buckets when you run a sensitive data discovery job. In addition, you can now view the estimated cost of a job when you create a job.	September 3, 2020
New content	Added information about configuring, running, and managing sensitive data discovery jobs.	August 31, 2020
New functionality	Managed data identifiers can now detect certain types of personally identifiable information for Brazil.	July 31, 2020
Updated content	Added information about the supported syntax for regular expressions in custom data identifiers.	July 30, 2020
Updated content	Added keyword requirements for managed data identifiers, and increased the quota for the number of findings that each sensitive data discovery job can produce.	July 17, 2020

New content	Added information about using Amazon EventBridge and AWS Security Hub to monitor and process findings. This includes the EventBridge event schema for findings and event examples for policy and sensitive data findings.	June 22, 2020
New content	Added information about analyzing and suppressing findings.	June 17, 2020
New content	Added instructions for configuring Macie to store detailed discovery results in an S3 bucket.	June 2, 2020
New content	Added information about the types of sensitive data that Macie can detect, and encryption requirements for detecting sensitive data in Amazon S3 objects.	May 28, 2020
General availability	This is the initial public release of the *Amazon Macie User Guide*.	May 13, 2020