Amazon Macie

User Guide

Amazon Macie: User Guide

Table of Contents

What Is Amazon Macie?

Amazon Macie is a security service that uses machine learning to automatically discover, classify, and protect sensitive data in AWS. Macie recognizes sensitive data such as personally identifiable information (PII) or intellectual property. It provides you with dashboards and alerts that give visibility into how this data is being accessed or moved.

Important
Currently, Macie is supported in the following Regions:

- US East (N. Virginia)
- US West (Oregon)

Features of Amazon Macie

Data Discovery and Classification

Amazon Macie enables you to identify business-critical data and analyze access patterns and user behavior as follows:

- Continuously monitor new data in your AWS environment
- Use artificial intelligence to understand access patterns of historical data
- Automatically access user activity, applications, and service accounts
- Use natural language processing (NLP) methods to understand data
- Intelligently and accurately assign business value to data and prioritize business-critical data based on your unique organization
- Create your own security alerts and custom policy definitions

Data Security

Amazon Macie enables you to be proactive with security compliance and achieve preventive security as follows:

- Identify and protect various data types, including PII, PHI, regulatory documents, API keys, and secret keys
- Verify compliance with automated logs that allow for instant auditing
- Identify changes to policies and access control lists
- Observe changes in user behavior and receive actionable alerts
- Receive notifications when data and account credentials leave protected zones
- Detect when large quantities of business-critical documents are shared internally and externally

Pricing for Macie

Pricing in Macie is based on the content sources classified or processed. For detailed information about Macie pricing, see Amazon Macie Pricing.

Accessing Macie

You can work with Macie in any of the following ways:

Macie console

Sign in to the AWS Management Console and open the Macie console at https://us-east-1.redirection.macie.aws.amazon.com/.

The console is a browser-based interface to access and use Macie.

Concepts and Terminology

As you get started with Amazon Macie, you can benefit from learning about its key concepts.

Account

A standard AWS account that contains your AWS resources. When you sign up for Amazon Web Services (AWS), your account is automatically signed up for all services in AWS, including Macie. The account that you use to sign in to AWS at the time when you first enable Macie is designated as the *master* account.

You can also integrate other accounts with Macie. These other accounts are called *member* accounts.

> **Note**
> No users from the member accounts are granted access to the Macie console. Only the master account users have access to the Macie console, where they can configure Macie and monitor and protect the resources in both master and member accounts.

Alert

A notification about a potential security issue that Macie discovers. Alerts appear on the Macie console and provide a comprehensive narrative about all activity that occurred over the last 24 hours.

Macie provides the following types of alerts:

- **Basic alerts** – Alerts that are generated by the security checks that Macie performs. There are two types of basic alerts in Macie:
  - Managed (curated by Macie) basic alerts that you can't modify. You can only enable or disable the existing managed basic alerts.
  - Custom basic alerts that you can create and modify to your exact specifications.
- **Predictive alerts** – Automatic alerts based on activity in your AWS infrastructure that deviates from the established normal activity baseline. More specifically, Macie continuously monitors IAM user and role activity in your AWS infrastructure and builds a model of the normal behavior. It then looks for deviations from that normal baseline, and when it detects such activity, it generates automatic predictive alerts. For example, a user uploading or downloading a large number of S3 objects in a day might trigger an alert if that user typically downloads one or two S3 objects ina week.

For more information about alerts, including alert categories and details about the contents of Macie alerts, see Amazon Macie Alerts (p. 41).

Data source

The origin or location of a set of data. To classify and protect your data, Macie analyzes and processes information from the following data sources:

AWS CloudTrail event logs, including Amazon S3 object-level API activity

AWS CloudTrail provides you with a history of AWS API calls for your account, including API calls made using the AWS Management Console, the AWS SDKs, the command line tools, and higher-level AWS services. AWS CloudTrail also enables you to identify which users and accounts called AWS APIs for services that support CloudTrail, the source IP address that the calls were made from, and when the calls occurred. For more information, see What Is AWS CloudTrail?

For data classification purposes, Macie uses the ability in CloudTrail to capture object-level API activity on S3 objects (data events). For more information, see Logging Data and Management Events for Trails.

Amazon S3

In this release, Macie analyzes and processes data stored in the Amazon S3 buckets. You can select the S3 buckets that contain objects that you want Macie to classify and monitor.

Amazon Simple Storage Service (Amazon S3) is storage for the Internet. Amazon S3 stores data as objects in buckets. An object consists of a file and optionally any metadata that describes that file. To store an object in Amazon S3, you upload the file that you want to store to a bucket. Buckets are the containers for objects. For more information, see Getting Started with Amazon Simple Storage Service.

User

In the context of Macie, a user is the AWS Identity and Access Management (IAM) identity that makes the request. Macie uses the CloudTrail `userIdentity` element to distinguish the following user types. For more information, see CloudTrail userIdentity Element.

- Root – The request was made with your account credentials.
- IAM user – The request was made with the credentials of an IAM user.
- Assumed role – The request was made with temporary security credentials that were obtained with a role via a call to the AWS Security Token Service (AWS STS) `AssumeRole` API operation.
- Federated user – The request was made with temporary security credentials that were obtained via a call to the AWS STS `GetFederationToken` API operation.
- AWS account – The request was made by another account.
- AWS service – The request was made by an account that belongs to an AWS service.

When specifying a user in the Macie console, you must use a special Macie format called `macieUniqueId`. Examples of specifying a user include searching for a user in the **Users** tab, constructing a query in the **Research** tab, and whitelisting a user in a basic alert with the index of **CloudTrail data**. The `macieUniqueId` is a combination of the IAM `UserIdentity` element and the `recipientAccountId`. For more information, see the preceding list of `UserIdentity` elements and the definition of `recipientAccountId` in the CloudTrail Record Contents. The following examples list various structures of `macieUniqueId`, depending on the user identity type:

- `123456789012:root`
- `123456789012:user/Bob`
- `123456789012:assumed-role/Accounting-Role/Mary`

For more detailed examples, see Analyzing Amazon Macie–Monitored Data by User Activity (p. 48).

Setting Up Amazon Macie

When you sign up for AWS, your AWS account is automatically signed up for all services in AWS, including Amazon Macie. If you don't have an account, use the following procedure to create one.

To sign up for AWS

1. Open https://aws.amazon.com/, and then choose **Create an AWS Account**.

   Note
   If you previously signed in to the AWS Management Console using AWS account root user credentials, choose **Sign in to a different account**. If you previously signed in to the console using IAM credentials, choose **Sign-in using root account credentials**. Then choose **Create a new AWS account**.

2. Follow the online instructions.

   Part of the sign-up procedure involves receiving a phone call and entering a verification code using the phone keypad.

Topics

Step 1: Enable Macie

When you launch the Macie console for the first time, choose **Get Started** and complete the following procedure to enable Macie.

Important
The AWS account that you use to enable Macie is automatically designated as your master account. For more information, see Concepts and Terminology (p. 3).

To enable Amazon Macie

1. The IAM identity (user, role, group) that you use to enable Macie must have the required permissions. To grant the permissions required to enable Macie, attach the `AmazonMacieFullAccess` managed policy to this IAM user, group, or role. For more information, see Managed (Predefined) Policies for Macie (p. 87).
2. Use the credentials of the IAM identity from step 1 to sign in to the Macie console. When you open the Macie console for the first time, choose **Get Started**.
3. On the **Enable Amazon Macie** page, verify Region preferences by reviewing the value in the dropdown menu under the **Region** section.

   Note
   The region that you're signed in to is automatically selected.

4. Choose **Enable Macie**.

Note the following about enabling Macie:

- Macie is assigned a service-linked role called `AWSServiceRoleForAmazonMacie`. This service-linked role includes the permissions and trust policy that Macie requires to discover, classify, and protect sensitive data in AWS on your behalf and to generate alerts about potential security issues. To view the details of `AWSServiceRoleForAmazonMacie`, on the **Enable Amazon Macie** page, choose **View service role permissions**. For more information, see Using Service-Linked Roles for Amazon Macie (p. 7). For more information about service-linked roles, see Using Service-Linked Roles.
- After you enable Macie, it immediately begins pulling and analyzing independent streams of data from AWS CloudTrail to generate alerts. Because Macie consumes this data only to determine if there are potential security issues, Macie doesn't manage CloudTrail for you or make its events and logs available to you. If you have enabled CloudTrail independent of Macie, you continue to have the option to configure its settings through the CloudTrail console or APIs. For more information, see What Is AWS CloudTrail?
- You can disable Macie at any time to stop it from processing and analyzing CloudTrail events. For more information, see Disabling Amazon Macie and Deleting Collected Metadata (p. 88).

Step 2: Integrate Amazon S3 with Macie

To classify and protect your data, Macie analyzes and processes information from CloudTrail and Amazon S3. Enabling CloudTrail in your account is required to enable Macie. Integrating S3 with Macie (in other words, initially specifying one or more S3 buckets for Macie to monitor) is not required to enable Macie. However, we strongly recommend that you integrate with Amazon S3 as part of setting up Macie and specify at least one S3 bucket that contains objects that you want Macie to classify and monitor. For more information and details about how Macie classifies your data, see Classifying Data with Amazon Macie (p. 14).

When you integrate with Amazon S3, Macie creates a trail and a bucket to store the logs about the Amazon S3 object-level API activity (data events) that it will analyze, along with other CloudTrail logs that it processes.

Important
Macie has a default limit on the amount of data that it can classify in an account. Once this data limit is reached, Macie stops classifying the data in this account. The default data classification limit is 3 TB. You can contact AWS Support and request an increase to the default limit.

You can use the following procedure to integrate with Amazon S3 as part of setting up Macie:

1. Log in to AWS with the credentials of the account that is serving as your Macie master account.
2. On the Macie console's **Integrations** tab, choose the **S3 Resources** tab.
3. On the **S3 Resources** tab, select the account ID (master or member) in the **Select an account** dropdown and then choose **Select**.
4. On the **Integrate S3 resources with Macie** page, choose either **Edit** to edit the buckets/prefixes that are already integrated with Macie or **Add** to select new buckets/prefixes to integrate with Macie.

   Note
   You can select up to 250 S3 buckets and prefixes. You can only select S3 buckets in your current AWS Region.
5. Next, you must configure the methods for classifying your new and existing S3 objects. The continuous classification method is applied to new objects that are added or updated after your buckets selection is complete. Continuous classification of new data enables S3 object-level logging for all selected buckets and prefixes. The one-time classification method is applied only once to all of the existing objects in the selected S3 buckets/prefixes.

Important
In the current Macie release, you can only select **Full** as the setting for the continuous classification (classifying new objects), and you can select **Full** or **None** as the settings for the one-time classification method (classifying all existing objects).

Important
If you enable one-time classification, the following values are displayed for each selected bucket/prefix:

- **Total objects** – Total number of objects in this S3 bucket/prefix.
- **Processed estimate** – Total size of the data from this bucket that Macie will actually classify.
- **Cost estimate** – Cost estimate for classifying objects in this bucket.

Macie also provides you with several totals for all of the selected buckets:

- **Total size** – Total size of the data within all of the selected buckets.
- **Total number of objects** – Total number of objects in all of the selected S3 buckets.
- **Processed estimate** – Total size of the data from all of the selected buckets that Macie will actually classify.
- **Total cost estimate** – Total content classification cost estimate for all selected buckets.

The **Total cost estimate** for each bucket is based on the **Processed estimate** value for that bucket and the general Macie pricing of $5 per GB processed by the content classification engine. The total cost estimates are provided only for S3 buckets, not for prefixes. For more information, see Amazon Macie Pricing.
The **Processed estimate** value for each bucket is calculated as follows:

- If an object's size is less than 1 KB, 1 KB is added to the **Processed estimate** value. Otherwise, the object's actual size is added to the **Processed estimate** value.
- If the object's size is greater than 20 MB, 20 MB is added to the **Processed estimate** value. Otherwise, the object's actual size is added to the **Processed estimate** value.
- For objects in Amazon S3 Glacier vaults, 0 is added to the **Processed estimate** value.

It's possible for the **Processed estimate** value of an S3 bucket to be higher than the **Total size** value.
The one-time classification cost estimates are only calculated per S3 buckets and *not* per S3 bucket prefixes. If you select an S3 bucket prefix, the cost estimate for the entire S3 bucket is included in the total cost estimate summary for the selected resources. If you select multiple prefixes of the same S3 bucket, the cost estimate for this entire S3 bucket is included only once in the total cost estimate summary for the selected resources.

6. When you have finished your selections, choose **Review**.
7. When you have finished reviewing your selections, choose **Start classification**.

Using Service-Linked Roles for Amazon Macie

Amazon Macie uses AWS Identity and Access Management (IAM) service-linked roles. A service-linked role is a unique type of IAM role that is linked directly to Macie. Service-linked roles are predefined by Macie and include all the permissions that Macie requires to call other AWS services on your behalf.

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 its service-linked role, and unless the

permissions are 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.

You can delete the Macie service-linked role only after first disabling Macie. This protects your Macie resources because you can't inadvertently remove permission to access them.

For information about other services that support service-linked roles, see AWS Services That Work with IAM in the *IAM User Guide* and look for the services that have **Yes** in the **Service-Linked Role** column. Choose a **Yes** with a link to view the service-linked role documentation for that service.

Service-Linked Role Permissions for Macie

Macie uses the service-linked role named `AWSServiceRoleForAmazonMacie`. It allows Amazon Macie to discover, classify, and protect sensitive data in AWS on your behalf.

The `AWSServiceRoleForAmazonMacie` service-linked role trusts the following services to assume the role:

- `macie.amazonaws.com`

The role permissions policy allows Macie to complete the following actions on the specified resources:

- Action: `iam:CreateServiceLinkedRole`
- Resources: `arn:aws:iam::*:role/aws-service-role/macie.amazonaws.com/AWSServiceRoleForAmazonMacie`

You must configure permissions to allow an IAM entity (such as a user, group, or role) to create, edit, or delete a service-linked role. For the `AWSServiceRoleForAmazonMacie` service-linked role to be successfully created, the IAM identity that you use Macie with must have the required permissions. To grant the required permissions, attach the `AmazonMacieFullAccess` managed policy to this IAM user, group, or role. For more information, see Service-Linked Role Permissions in the *IAM User Guide*.

Creating a Service-Linked Role for Macie

- **For the master Macie account**, the `AWSServiceRoleForAmazonMacie` service-linked role is automatically created when you enable Macie for the first time or enable Macie in a supported Region where you previously didn't have it enabled. You can also create the `AWSServiceRoleForAmazonMacie` service-linked role manually for the master account using the IAM console, the IAM CLI, or the IAM API.
- **For member Macie accounts**, the `AWSServiceRoleForAmazonMacie` service-linked role is automatically created when the master account associates a member account with Macie. You can also create `AWSServiceRoleForAmazonMacie` for member accounts manually using the IAM console, the IAM CLI, or the IAM API.

  Important
  The service-linked role that is created for the master Macie account doesn't apply to the member Macie accounts.

For more information about creating the role manually, see Creating a Service-Linked Role in the *IAM User Guide*.

Important
If you were using the Macie service before June 21, 2018, when it began supporting service-linked roles, the IAM roles that grant Macie access to call other AWS services on

your behalf already exist in your AWS account (Macie master or member). These roles are `AmazonMacieServiceRole` and `AmazonMacieSetupRole`. They were created when you launched either the Macie AWS CloudFormation template for a master account or the Macie AWS CloudFormation template for a member account as part of setting up Macie.
The newly created service-linked role replaces these previously created IAM roles (in master and member accounts).

Note
These previously created IAM roles aren't deleted. They remain intact, but they're no longer used to grant Macie access to call other AWS services on your behalf. You can use the IAM console to manage or delete these IAM roles.

Editing a Service-Linked Role for Macie

Macie doesn't allow you to edit the `AWSServiceRoleForAmazonMacie` service-linked role. After you create a service-linked role, 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 using IAM. For more information, see Editing a Service-Linked Role in the *IAM User Guide*.

Deleting a Service-Linked Role for Macie

If you no longer need to use a feature or service that requires a service-linked role, we recommend that you delete that role. That way, you don't have an unused entity that isn't actively monitored or maintained.

Important
For a master account, you must first disable Macie to delete the `AWSServiceRoleForAmazonMacie`.
For Macie member accounts, to delete the service-linked role used in a member account, the master Macie account must first disassociate this member account from Macie.
If the Macie service isn't disabled when you try to delete the service-linked role, the deletion fails. For more information, see Disabling Amazon Macie and Deleting Collected Metadata (p. 88).

When you disable Macie, the `AWSServiceRoleForAmazonMacie` is *not* automatically deleted. If you enable Macie again, it starts using the existing `AWSServiceRoleForAmazonMacie`.

To manually delete the service-linked role using IAM

Use the IAM console, the IAM CLI, or the IAM API to delete the `AWSServiceRoleForAmazonMacie` service-linked role. For more information, see Deleting a Service-Linked Role in the *IAM User Guide*.

Integrating Member Accounts and Amazon S3 with Amazon Macie

You can use the Macie console's **Integrations** tab to integrate member accounts with Macie and to integrate Amazon S3 with Macie for both your master account and member accounts. For more information about the master and member accounts, see Concepts and Terminology (p. 3).

Topics

Integrating Member Accounts with Macie

When you integrate member accounts with Macie, you're enabling Macie to monitor resources and activity in these member accounts.

To integrate member accounts with Macie

1. Log in to AWS with the credentials of the AWS account that you want to integrate with Macie as a member account.
2. **Important**
   This is a required step if you're adding an account as a Macie member account for the first time.
   You can skip this step if you're readding an account as a Macie member account after disassociating it from Macie.

   Create the IAM role called `AmazonMacieHandshakeRole` that grants this account the required permissions to be successfully integrated with Macie. You can create this role by launching the AWS CloudFormation stack templates found at the URLs listed below. You need to create this role only once for use in all Regions. For more information about AWS CloudFormation and AWS CloudFormation stacks, see What is AWS CloudFormation? and Working with Stacks.

   Important
   Make sure to specify the master account ID when running the following stack templates:

   - US East (N. Virginia): Macie AWS CloudFormation template for a member account
   - US West (Oregon): Macie AWS CloudFormation template for a member account
3. Log in to AWS with the Macie master account, navigate to the Macie console, and choose the **Integrations** tab.
4. To integrate a member account, choose the **+** icon next to **Member accounts**.
5. In the **Add member AWS account(s)** pop-up window, enter one or more account IDs. Separate multiple account numbers with commas. Choose **Add accounts**.

   Important
   Once Macie is enabled in this member account, Macie is assigned a service-linked role called `AWSServiceRoleForAmazonMacie`. This service-linked role includes the permissions and the

trust policy that Macie requires to discover, classify, and protect sensitive data in AWS in this account and to generate alerts about potential security issues. The following are the details of `AWSServiceRoleForAmazonMacie`.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Resource": "*",
            "Action": [
                "cloudtrail:DescribeTrails",
                "cloudtrail:GetEventSelectors",
                "cloudtrail:GetTrailStatus",
                "cloudtrail:ListTags",
                "cloudtrail:LookupEvents",
                "iam:ListAccountAliases",
                "s3:Get*",
                "s3:List*"
            ]
        },
        {
            "Effect": "Allow",
            "Resource": "arn:aws:cloudtrail:*:*:trail/AWSMacieTrail-DO-NOT-EDIT",
            "Action": [
                "cloudtrail:CreateTrail",
                "cloudtrail:StartLogging",
                "cloudtrail:StopLogging",
                "cloudtrail:UpdateTrail",
                "cloudtrail:DeleteTrail",
                "cloudtrail:PutEventSelectors"
            ]
        },
        {
            "Effect": "Allow",
            "Resource": [
                "arn:aws:s3:::awsmacie-*",
                "arn:aws:s3:::awsmacietrail-*",
                "arn:aws:s3:::*-awsmacietrail-*"
            ],
            "Action": [
                "s3:CreateBucket",
                "s3:DeleteBucket",
                "s3:DeleteBucketPolicy",
                "s3:DeleteBucketWebsite",
                "s3:DeleteObject",
                "s3:DeleteObjectTagging",
                "s3:DeleteObjectVersion",
                "s3:DeleteObjectVersionTagging",
                "s3:DeleteReplicationConfiguration",
                "s3:PutBucketPolicy"
            ]
        }
    ]
}
```

For more information, see Using Service-Linked Roles for Amazon Macie (p. 7). For more information about service-linked roles, see Using Service-Linked Roles.

If you disable Macie, the service no longer has access to the resources in your member account. However, after you disable Macie, the `AmazonMacieHandshakeRole` IAM role that you created for this member account in step 2 and the `AWSServiceRoleForAmazonMacie` service-linked role that was automatically

created for this member account when it was integrated with Macie remain intact. These existing roles are used again if you decide to reenable Macie for this member account. To reenable Macie for a member account, you must use the previous steps to integrate this member account with Macie.

Specifying Data for Macie to Monitor

You can use the **S3 Resources** tab on the **Integrations** tab to specify the S3 buckets and prefixes that contain the data that you want Macie to monitor.

Important
Currently, Macie can only monitor objects stored in Amazon S3.

Important
Macie has a default limit on the amount of data that it can classify in an account. Once this data limit is reached, Macie stops classifying the data in this account. The default data classification limit is 3 TB. You can contact AWS Support and request an increase to the default limit.

You can also specify S3 buckets and prefixes for Macie to monitor during the initial Macie setup. For more information and instructions, see Setting Up Amazon Macie (p. 5).

1. Log in to AWS with the credentials of the account that is your Macie master account.
2. In the Macie console's **Integrations** tab, choose the **S3 Resources** tab.
3. In the **S3 Resources** tab, select the account ID (master or member) in the **Select an account** dropdown and then choose **Select**.
4. On the **Integrate S3 resources with Macie** page, choose either **Edit** to edit the buckets/prefixes that are already integrated with Macie or **Add** to select new buckets/prefixes to integrate with Macie.

   Note
   You can select up to 250 S3 buckets and prefixes. You can only select S3 buckets in your current AWS Region.
5. Configure the methods for classifying your new and existing S3 objects. The continuous classification method is applied to new objects that are added or updated after your buckets selection is complete. Continuous classification of new data enables S3 object-level logging for all selected buckets and prefixes. The one-time classification method is applied only once to all of the existing objects in the selected S3 buckets/prefixes.

   Important
   In the current Macie release, you can only select **Full** as the setting for the continuous classification (classifying new objects), and you can select **Full** or **None** as the settings for the one-time classification method (classifying all existing objects).

   Important
   If you enable one-time classification, note the following values that are displayed for each selected bucket/prefix:

   - Total objects – Total number of objects in this S3 bucket/prefix.
   - Processed estimate – Total size of the data from this bucket that Macie will actually classify.
   - Cost estimate – Cost estimate for classifying objects in this bucket.

   Macie also provides you with several totals for all of the selected buckets:

   - **Total size** – Total size of the data within all of the selected buckets.
   - **Total number of objects** – Total number of objects in all of the selected S3 buckets.
   - **Processed estimate** – Total size of the data from all of the selected buckets that Macie will actually classify.

- **Total cost estimate** – Total content classification cost estimate for all selected buckets.

  The **Total cost estimate** for each bucket is based on the **Processed estimate** value for that bucket and the general Macie pricing of $5 per GB processed by the content classification engine. The total cost estimates are provided only for S3 buckets, not for prefixes. For more information, see Amazon Macie Pricing.
  The **Processed estimate** value for each bucket is calculated as follows:

  - If an object's size is less than 1 KB, 1 KB is added to the **Processed estimate** value. Otherwise, the object's actual size is added to the **Processed estimate** value.
  - If the object's size is greater than 20 MB, 20 MB is added to the **Processed estimate** value. Otherwise, the object's actual size is added to the **Processed estimate** value.
  - For objects in Amazon S3 Glacier vaults, 0 is added to the **Processed estimate** value.

  It's possible for the **Processed estimate** value of an S3 bucket to be higher than the **Total size** value.
  The one-time classification cost estimates are only calculated per S3 buckets and *not* per S3 bucket prefixes. If you select an S3 bucket prefix, the cost estimate for the entire S3 bucket is included in the total cost estimate summary for the selected resources. If you select multiple prefixes of the same S3 bucket, the cost estimate for this entire S3 bucket is included only once in the total cost estimate summary for the selected resources.

6. When you have finished your selections, choose **Review**.
7. When you have finished reviewing your selections, choose **Start classification**.

Encrypted Objects

If objects stored in your Amazon S3 buckets are encrypted, Macie might not be able to read and classify those objects for the following reasons:

- If your Amazon S3 objects are encrypted using Amazon S3–managed encryption keys (SSE-S3), Macie can read and classify the objects using the roles created during the setup process.
- If your Amazon S3 objects are encrypted using AWS KMS–managed keys (SSE-KMS), Macie can read and classify the objects only if you add the `AWSMacieServiceCustomerServiceRole` IAM role or the `AWSServiceRoleForAmazonMacie` service-linked role as a key user for the KMS customer master key (CMK). If you don't add either of these roles as a key user for the KMS CMK, Macie can't read and classify the objects. However, Macie still stores metadata on the object, including which KMS CMK was used to protect the object.
- If your Amazon S3 objects are encrypted using client-side encryption, Macie can't read and classify the objects, but still stores metadata on the object.

Classifying Data with Amazon Macie

Macie can help you classify your sensitive and business-critical data stored in the cloud. Currently, Macie analyzes and processes data stored in Amazon S3 buckets. To classify your data, Macie also uses the ability in AWS CloudTrail to capture object-level API activity on S3 objects (data events). However, Macie monitors CloudTrail data events only if you specify at least one S3 bucket for Macie to monitor.

Once you specify the S3 bucket or buckets for Macie to monitor, you enable Macie to continuously monitor and discover new data as it enters your AWS infrastructure. For more information on how to specify S3 buckets for Macie to monitor, see Specifying Data for Macie to Monitor (p. 12).

Note
Macie's content classification engine processes up to the first 20 MB of an S3 object.

If you specify S3 buckets that include files of a format that isn't supported in Macie, Macie doesn't classify them, and your Macie usage charges don't include any costs for this content. Your Macie usage charges include only the costs for the content that Macie processes. For example, Macie can't extract text from .wav files (images or movies); therefore, it doesn't process that content, and you're not charged for it.

Topics

Supported Compression and Archive File Formats

Currently, Macie supports the following compression and archive file formats:

- BZIP
- GZIP
- LZO
- RAR
- SNAPPY
- AR
- CPIO
- Unix dump
- TAR
- zip
- XZ
- Pack200
- BZIP2
- 7z

- ARJ
- LZMA
- DEFLATE
- Brotli

Content Type

Once Macie begins monitoring your data, it uses several automatic content classification methods to identify and prioritize your sensitive and critical data and to accurately assign business value to your data. One of these methods is classifying by content type.

To classify your data objects by content type, Macie uses an identifier that is embedded in the file header. Macie offers a set of managed (Macie-curated) content types, each with a designated risk level between 1 and 10.

Macie can assign only one content type to an object.

You can't modify existing or add new content types. You can enable or disable any existing content types, thus enabling or disabling Macie to assign these them to your objects during the classification process.

To view, enable, or disable content types

1. In the Macie console, navigate to the **Settings** page.
2. In the **Classify data** section, choose **Content types**.
3. Choose any of the listed managed content types to view its details.

   To enable or disable a content type, on its details page, use the **Enabled/Disabled** dropdown and choose **Save**.

The following list describes the complete list of content types that Macie can assign to your objects.

Name	Description
application/cap	WireShark or Tcpdump Packet Capture
application/epub+zip	application/epub
application/illustrator	Adobe Illustrator
application/java	Binary (Java)
application/java-archive	application/java-archive
application/java-serialized-object	application/java-serialized-object
application/java-vm	application/java-vm
application/javascript	application/javascript
application/json	JSON
application/msaccess	application/msaccess
application/msexcel	Microsoft Excel
application/msonenote	application/msonenote

application/mspowerpoint	Microsoft PowerPoint
application/msword	Microsoft Word
application/octet-stream	application/octet-stream
application/octet-stream+fon	application/octet-stream+fon
application/ogg	application/ogg
application/onenote	application/onenote
application/pdf	Adobe PDF
application/pgp	application/pgp
application/pgp-encrypted	application/pgp-encrypted
application/pgp-keys	PGP keys
application/pgp-signature	PGP signature
application/postscript	Adobe Postscript
application/rar	RAR compressed archive
application/rdf+xml	application/rdf+xml
application/rss+xml	application/rss+xml
application/rtf	application/rtf
application/tar	TAR archive
application/unknown	application/unknown
application/vnd.3gpp.pic-bw-small	application/vnd.3gpp.pic-bw-small
application/vnd.android.package-archive	Android Package
application/vnd.audiograph	application/vnd.audiograph
application/vnd.balsamiq.bmpr	Balsamiq Mockup
application/vnd.cups-ppd	application/vnd.cups-ppd
application/vnd.curl.car	application/vnd.curl.car
application/vnd.dvb.ait	application/vnd.dvb.ait
application/vnd.google-apps.document	Google Apps Document
application/vnd.google-apps.drawing	application/vnd.google-apps.drawing
application/vnd.google-apps.form	Google Apps Form
application/vnd.google-apps.map	Google Apps Map
application/vnd.google-apps.presentation	Google Apps Presentation
application/vnd.google-apps.script	Google Apps script
application/vnd.google-apps.spreadsheet	Google Apps Spreadsheet
application/vnd.google-earth.kmz	Google Earth KMZ

application/vnd.jcp.javame.midlet-rms	application/vnd.jcp.javame.midlet-rms
application/vnd.jgraph.mxfile	application/vnd.jgraph.mxfile
application/vnd.jgraph.mxfile.realtime	application/vnd.jgraph.mxfile.realtime
application/vnd.jgraph.mxfile.rtlegacy	application/vnd.jgraph.mxfile.rtlegacy
application/vnd.kde.kontour	application/vnd.kde.kontour
application/vnd.lotus-1-2-3	application/vnd.lotus-1-2-3
application/vnd.lotus-organizer	application/vnd.lotus-organizer
application/vnd.mozilla.xul+xml	application/vnd.mozilla.xul+xml
application/vnd.ms-excel	Excel
application/vnd.ms-excel.addin.macroEnabled.12	application/vnd.ms-excel.addin.macroEnabled.12
application/vnd.ms-excel.sheet.binary.macroEnabled.12	Microsoft Excel - Macro enabled
application/vnd.ms-excel.sheet.macroEnabled.12	Microsoft Excel - Macro enabled
application/vnd.ms-excel.sheet.macroenabled.12	application/vnd.ms-excel.sheet.macroenabled.12
application/vnd.ms-excel.template.macroenabled.12	application/vnd.ms-excel.template.macroenabled.12
application/vnd.ms-fontobject	application/vnd.ms-fontobject
application/vnd.ms-htmlhelp	application/vnd.ms-htmlhelp
application/vnd.ms-officetheme	application/vnd.ms-officetheme
application/vnd.ms-package.relationships+xml	application/vnd.ms-package.relationships+xml
application/vnd.ms-pki.seccat	Microsoft Exchange Server Certificate Store
application/vnd.ms-powerpoint	Microsoft PowerPoint
application/vnd.ms-powerpoint.presentation.macroEnabled.12	application/vnd.ms-powerpoint.presentation.macroEnabled.12
application/vnd.ms-powerpoint.slideshow.macroEnabled.12	application/vnd.ms-powerpoint.slideshow.macroEnabled.12
application/vnd.ms-powerpointtd	Microsoft PowerPoint
application/vnd.ms-project	application/vnd.ms-project
application/vnd.ms-publisher	application/vnd.ms-publisher
application/vnd.ms-word.document.macroEnabled.12	Microsoft Word - Macro enabled
application/vnd.ms-xpsdocument	application/vnd.ms-xpsdocument
application/vnd.oasis.opendocument.chart	application/vnd.oasis.opendocument.chart

application/vnd.oasis.opendocument.graphics	application/vnd.oasis.opendocument.graphics
application/vnd.oasis.opendocument.presentation	Presentation
application/vnd.oasis.opendocument.spreadsheet	Spreadsheet
application/vnd.oasis.opendocument.text	Open Document Text
application/vnd.openxmlformats-officedocument.presentationml.presentation	Microsoft PowerPoint
application/vnd.openxmlformats-officedocument.presentationml.slide	Microsoft Powerpoint
application/vnd.openxmlformats-officedocument.presentationml.slideshow	Microsoft Powerpoint
application/vnd.openxmlformats-officedocument.presentationml.template	application/vnd.openxmlformats-officedocument.presentationml.template
application/vnd.openxmlformats-officedocument.spreadsheetml.sheet	Microsoft Excel
application/vnd.openxmlformats-officedocument.spreadsheetml.template	application/vnd.openxmlformats-officedocument.spreadsheetml.template
application/vnd.openxmlformats-officedocument.wordprocessingml.document	Microsoft Word
application/vnd.openxmlformats-officedocument.wordprocessingml.template	application/vnd.openxmlformats-officedocument.wordprocessingml.template
application/vnd.palm	application/vnd.palm
application/vnd.symbian.install	application/vnd.symbian.install
application/vnd.tcpdump.pcap	Wireshark or Tcpdump Packet Capture
application/vnd.visio	Microsoft Visio
application/vns.ms-outlook	Microsoft Outlook messages
application/x-7z-compressed	7zip compressed archive
application/x-adobeaamdetect	Adobe Application Manager
application/x-adobeexmandetect	application/x-adobeexmandetect
application/x-apple-diskimage	Apple disk image
application/x-bittorrent	application/x-bittorrent
application/x-bzip2	application/x-bzip2
application/x-cab	application/x-cab
application/x-cfs-compressed	application/x-cfs-compressed
application/x-coredump	application/x-coredump
application/x-couponprinterplugin	application/x-couponprinterplugin

application/x-dbm	application/x-dbm
application/x-dosexec	Executable
application/x-dvi	application/x-dvi
application/x-executable	Executable
application/x-fla	application/x-fla
application/x-font	application/x-font
application/x-font-otf	application/x-font-otf
application/x-font-ttf	application/x-font-ttf
application/x-font-type1	application/x-font-type1
application/x-font-woff	application/x-font-woff
application/x-freemind	application/x-freemind
application/x-gtar	GNU tar compressed archive
application/x-gzip	GNU Zip compressed archive
application/x-iso9660-image	application/x-iso9660-image
application/x-iwork-keynote-sffkey	application/x-iwork-keynote-sffkey
application/x-iwork-numbers-sffnumbers	application/x-iwork-numbers-sffnumbers
application/x-iwork-pages-sffpages	application/x-iwork-pages-sffpages
application/x-javascript	application/x-javascript
application/x-maker	application/x-maker
application/x-mobipocket-ebook	application/x-mobipocket-ebook
application/x-ms-shortcut	application/x-ms-shortcut
application/x-ms-wmz	application/x-ms-wmz
application/x-msdos-program	Microsoft Windows Application
application/x-msi	application/x-msi
application/x-msmetafile	application/x-msmetafile
application/x-mspublisher	application/x-mspublisher
application/x-nawk	application/x-nawk
application/x-ns-proxy-autoconfig	application/x-ns-proxy-autoconfig
application/x-object	application/x-object
application/x-perl	Perl Source Code
application/x-pkcs12	PKI Certificate
application/x-pkcs7-crl	PKI Files
application/x-python-code	Source Code (python)

application/x-rar-compressed	RAR compressed archive
application/x-redhat-package-manager	application/x-redhat-package-manager
application/x-sas	Statistical Analysis
application/x-sharedlib	application/x-sharedlib
application/x-shellscript	Shell Script
application/x-shockwave-flash	application/x-shockwave-flash
application/x-silverlight-app	application/x-silverlight-app
application/x-stuffit	Stuffit compressed archive
application/x-subrip	application/x-subrip
application/x-tar	TAR archive
application/x-tex-tfm	Apache Font
application/x-texinfo	application/x-texinfo
application/x-troff-man	application/x-troff-man
application/x-wais-source	application/x-wais-source
application/x-x509-ca-cert	application/x-x509-ca-cert
application/x-xcf	application/x-xcf
application/x-xfig	application/x-xfig
application/x-xpinstall	application/x-xpinstall
application/x-zip	Zip compressed archive
application/xhtml+xml	application/xhtml+xml
application/xmind	application/xmind
application/xml	XML Text
application/xv+xml	application/xv+xml
application/zip	Zip compressed archive
binary/octet-stream	binary/octet-stream
chemical/x-cache	chemical/x-cache
chemical/x-cerius	chemical/x-cerius
chemical/x-gamess-input	chemical/x-gamess-input
chemical/x-genbank	chemical/x-genbank
chemical/x-mdl-sdfile	chemical/x-mdl-sdfile
chemical/x-pdb	Protein Databank chemical/x-pdb
chemical/x-rosdal	chemical/x-rosdal
message/rfc822	message/rfc822

text/cache-manifest	text/cache-manifest
text/calendar	text/calendar
text/css	text/css
text/csv	Comma Separated Values
text/html	text/html
text/json	JavaScript Object Notation
text/plain	Plain Text
text/rtf	text/rtf
text/tab-separated-values	Tab separated values
text/texmacs	text/texmacs
text/vnd.graphviz	text/vnd.graphviz
text/x-asm	Source Code (Assembly)
text/x-bibtex	text/x-bibtex
text/x-c	Source Code (c)
text/x-c++hdr	Source Code (C++ headers)
text/x-c++src	Source Code (c++)
text/x-chdr	Source Code (C headers)
text/x-component	text/x-component
text/x-csh	Source Code (C shell)
text/x-csharp	Source Code (C#)
text/x-csrc	Source Code (C)
text/x-diff	text/x-diff
text/x-dsrc	text/x-dsrc
text/x-java	Source Code (Java)
text/x-java-source	Source Code (Java)
text/x-markdown	text/x-markdown
text/x-nfo	text/x-nfo
text/x-objcsrc	Source Code (Objective-C)
text/x-pascal	Source Code (Pascal)
text/x-perl	Source Code (Perl)
text/x-python	Source Code (Python)
text/x-sfv	text/x-sfv
text/x-sh	Source Code (x-sh)

text/x-sql	Source Code (SQL)
text/x-tex	text/x-tex
text/x-url	text/x-url
text/x-vcard	text/x-vcard
text/xml	XML Text

File Extension

Once Macie begins monitoring your data, it uses several automatic content classification methods to identify and prioritize your sensitive and critical data and to accurately assign business value to your data. One of these methods is classifying by file extension.

Macie can also classify your objects by their file extensions. Macie offers a set of managed file extensions, each with a designated risk level between 1 and 10.

Macie can assign only one file extension to an object.

You can't modify existing or add new file extensions. You can enable or disable any existing file extensions, thus enabling or disabling Macie to assign them to your objects during the classification process.

To view, enable, or disable file extensions

1. In the Macie console, navigate to the **Settings** page.
2. In the **Classify data** section, choose **File extensions**.
3. Choose any of the listed managed file extensions to view its details.

   To enable or disable a file extension, on its details page, use the **Enabled/Disabled** dropdown and choose **Save**.

The following is the complete list of file extensions that Macie can assign to your objects during classification.

Name	Description
7z	7-Zip compressed file
abc	SolidWorks CAD
accdb	Microsoft Access database
apk	Application installable on Android
bat	Batch file
bin	Compressed archive. Readable by Java. Extractable by 7-zip
bz2	Bzip2 compressed archive
bzip2	Bzip2 compressed archive
c	C source code
c#	C# source code

cab	Microsoft cabinet. Extractable via ZIP
cc	C++ source code
cer	PKI certificate
cpp	C++ source code
csv	Comma Separated Values
cxx	C++ source code
dbf	dBase database
dbx	Microsoft Outlook Express
deb	Debian Linux install package
dmg	Apple OS X Application Installer
doc	Microsoft Word
docx	Microsoft Word
dot	Microsoft Word
dotx	Microsoft Word
dwg	AutoDesk CAD
dxf	AutoCAD
eml	MIME email
emlx	Apple Mail email message
exe	Microsoft Windows PE Executable
gpg	PGP certificate
gz	GNU Zip compressed archive
gzip	GNU Zip compressed archive
html	Hyper Text Markup Language
iwa	Apple iWork document archive file
jar	Java source code archive
java	Java source code
json	Java Script Object Notation Values (JSON)
key	Apple Keynote Presentation
keynote	Apple Keynote Presentation
lua	Lua source code
mdb	Microsoft Access database
msg	Microsoft Outlook Message
msi	Microsoft Windows Application Installer

odp	OpenOffice.org OpenDocument presentation file
oos	OpenOffice.org spreadsheet file
p12	PKI certificate
pages	Apple Pages
pdf	Adobe PDF
perl	Perl source code
pgp	PGP certificate
pl	Perl source code
pot	Microsoft PowerPoint
pps	Microsoft PowerPoint
ppt	Microsoft PowerPoint
pptx	Microsoft PowerPoint
pst	Microsoft Outlook
py	Python source code
rar	RAR archive. Extractable by 7-zip
rtf	Rich Text Format
sdp	OpenOffice.org presentation file
sdw	OpenOffice.org text document file
sldasm	SolidWorks CAD
slddrw	SolidWorks CAD
sldprt	SolidWorks CAD
sql	Structured Query Language
sxi	OpenOffice.org presentation file
sxw	OpenOffice.org Writer document file
tar.gz	GNU Zip compressed archive
tsv	Tab Separated Values
txt	Text Document
vdx	Microsoft Visio
vsd	Microsoft Visio
vss	Microsoft Visio
vst	Microsoft Visio
vsx	Microsoft Visio
vtw	Microsoft Visio

vtx	Microsoft Visio
xls	Microsoft Excel
xlsx	Microsoft Excel
xlw	Microsoft Excel
xml	Extensible Markup Language (XML)
xps	Open XML document specification
zip	ZIP compressed archive

Theme

Once Macie begins monitoring your data, it uses several automatic content classification methods to identify and prioritize your sensitive and critical data and to accurately assign business value to your data. One of these methods is classifying by theme.

Object classification by theme is based on keywords that Macie searches for as it examines the contents of data objects. Macie offers a set of managed themes, each with a designated risk level between 1 and 10.

Macie can assign one or more themes to an object.

You can't modify existing or add new themes. You can enable or disable any existing themes, thus enabling or disabling Macie to assign them to your objects during the classification process.

To view, enable, or disable themes

1. In the Macie console, navigate to the **Settings** page.
2. In the **Classify data** section, choose **Themes**.
3. Choose any of the listed managed themes to view its details.

   To enable or disable a theme, on its details page, use the **Enabled/Disabled** dropdown and then choose **Save**.

The following is the complete list of themes that Macie can assign to your objects during classification.

Theme title	Minimum keyword combinations
American Express Credit Card Keywords	1
Attorney Client Privileged	2
Audit Keywords	3
Banking Keywords	1
Big Data Frameworks	2
Cisco Analysis Keywords	1
Confidential Markings	2
Corporate Growth Keywords	3

Corporate Project Plan	3
Corporate Proposals	3
Credit Card Keywords	1
Encrypted Data Keywords	1
Financial Keywords	1
Hacker Keywords	2
Limit Distribution Markings	3
Mastercard Credit Card Keywords	1
Metasploit Framework Keywords	1
NMAP OS Fingerprinting	1
Network Scanner Keywords	1
Network Service Fingerprinting Keywords	1
Network Traffic Analysis Keywords	1
OS Backdoor Keywords	1
Offline Attacks Keywords	1
Online Attacks Keywords	1
Oracle DB Analysis Keywords	1
Password Keywords	2
Project Tracking Keywords	2
Proprietary Markings	2
Real-Time Processing Frameworks	2
Restricted Markings	2
SSL Forensic Analysis Keywords	1
Secret Markings	3
Sensitive Markings	3
Social Security Keywords	2
Stock Keywords	3
Taxpayer EIN Keywords	2
Tunneling Attacks Keywords	1
Unclassified Markings	2
VISA Credit Card Keywords	1
Vulnerability Assessment Keywords	2
Web Exploitation Tool Keywords	1

Web Vulnerability Scanner Keywords	1
pof OS Fingerprinting	2

Regex

Once Macie begins monitoring your data, it uses several automatic content classification methods to identify and prioritize your sensitive and critical data and to accurately assign business value to your data. One of these methods is classifying by regex.

Object classification by regex is based on specific data or data patterns that Macie searches for as it examines the contents of data objects. Macie offers a set of managed regexes, each with a designated risk level between 1 and 10.

Macie can assign one or more regexes to an object.

You can't modify existing or add new regexes. You can enable or disable any existing regexes, thus enabling or disabling Macie to assign them to your objects during the classification process.

To view, enable, or disable regexes

1. In the Macie console, navigate to the **Settings** page.
2. In the **Classify data** section, choose **Regex**.
3. Choose any of the listed managed regexes to view its details.

   To enable or disable a regex, on its details page, use the **Enabled/Disabled** dropdown and choose **Save**.

The following is the complete list of regexes that Macie can assign to your objects during classification.

Name	Classification
Arista network configuration	Regex
BBVA Compass Routing Number - California	Regex
Bank of America Routing Numbers - California	Regex
Box Links	Regex
CVE Number	Regex
California Drivers License	Regex
Chase Routing Numbers - California	Regex
Cisco Router Config	Regex
Citibank Routing Numbers - California	Regex
DSA Private Key	Regex
Dropbox Links	Regex
EC Private Key	Regex
Encrypted DSA Private Key	Regex

Encrypted EC Private Key	Regex
Encrypted Private Key	Regex
Encrypted PuTTY SSH DSA Key	Regex
Encrypted PuTTY SSH RSA Key	Regex
Encrypted RSA Private Key	Regex
Google Application Identifier	Regex
HIPAA PHI National Drug Code	Regex
Huawei config file	Regex
Individual Taxpayer Identification Numbers (ITIN)	Regex
John the Ripper	Regex
KeePass 1.x CSV Passwords	Regex
KeePass 1.x XML Passwords	Regex
Large number of US Phone Numbers	Regex
Large number of US Zip Codes	Regex
Lightweight Directory Access Protocol	Regex
Metasploit Module	Regex
MySQL database dump	Regex
MySQLite database dump	Regex
Network Proxy Auto-Config	Regex
Nmap Scan Report	Regex
PGP Header	Regex
PGP Private Key Block	Regex
PKCS7 Encrypted Data	Regex
Password etc passwd	Regex
Password etc shadow	Regex
PlainText Private Key	Regex
PuTTY SSH DSA Key	Regex
PuTTY SSH RSA Key	Regex
Public Key Cryptography System (PKCS)	Regex
Public encrypted key	Regex
RSA Private Key	Regex
SSL Certificate	Regex
SWIFT Codes	Regex

Samba Password config file	Regex
Simple Network Management Protocol Object Identifier	Regex
Slack 2FA Backup Codes	Regex
UK Drivers License Numbers	Regex
UK Passport Number	Regex
USBank Routing Numbers - California	Regex
United Bank Routing Number - California	Regex
Wells Fargo Routing Numbers - California	Regex
aws_access_key	Regex
aws_credentials_context	Regex
aws_secret_key	Regex
facebook_secret	Regex
github_key	Regex
google_two_factor_backup	Regex
heroku_key	Regex
microsoft_office_365_oauth_context	Regex
pgSQL Connection Information	Regex
slack_api_key	Regex
slack_api_token	Regex
ssh_dss_public	Regex
ssh_rsa_public	Regex

Personally Identifiable Information

Object classification by personally identifiable information (PII) is based on recognizing any personally identifiable artifacts based on industry standards such as NIST-80-122 and FIPS 199. Macie can recognize the following PII artifacts:

- Full names
- Mailing addresses
- Email addresses
- Credit card numbers
- IP addresses (IPv4 and IPv6)
- Drivers license IDs (USA)
- National identification numbers (USA)
- Birth dates

As part of PII object classification, Macie also assigns each matching object a PII impact of high, moderate, and low using the following criteria:

- High
  - >= 1 full name and credit card
  - >= 50 names or emails and any combination of other PII
- Moderate
  - >= 5 names or emails and any combination of other PII
- Low
  - 1–5 names or emails and any combination of PII
  - Any quantity of PII attributes above (without names or emails)

Support Vector Machine–Based Classifier

Another method that Macie uses to classify your S3 objects is the Support Vector Machine (SVM) classifier. It classifies content inside your S3 objects (text, token n-grams, and character n-grams) that Macie monitors and their metadata features (document length, extension, encoding, headers) to accurately classify documents based on content. This classifier, managed by Macie, was trained against a large corpus of training data of various types and has been optimized to support accurate detection of various content types, including source code, application logs, regulatory documents, and database backups. The classifier can also generalize its detections. For example, if it detected a new kind of source code that doesn't match any of the types of source code that it is trained to recognize, it can generalize the detection as being "source code."

> **Note**
> This data classification method isn't surfaced in the **Settings** page. Macie manages the following list of artifacts. You can't edit, enable, or disable them.

The SVM classifier in Macie is trained to detect the following content types:

- E-books
- Email
- Generic encryption keys
- Financial
  - SEC regulatory forms
- JSON
  - AWS CloudTrail logs
  - Jupyter notebooks
- Application logs
  - Apache format
  - Amazon S3 server logs
  - Linux syslog
- Database
  - MongoDB backup
  - MySQLbackup
  - MySQL script
- Source code
  - F#
  - VimL
  - ActionScript

  - Assembly
  - Bash
  - Batchfile
  - C
  - Clojure
  - Cobol
  - CoffeeScript
  - CUDA
  - Erlang
  - Fortran
  - Go
  - Haskell
  - Java
  - JavaScript
  - LISP
  - Lua
  - Matlab
  - ObjectiveC
  - Perl
  - PHP
  - PowerShell
  - Processing
  - Python
  - R
  - Ruby
  - Scala
  - Swift
  - VHDL
- Web languages
  - CSS
  - HTML
  - XML

Object Risk Level

Through the automatic classification methods previously described, an object that Macie monitors is assigned various risk levels based on each content type, file extension, theme, regex, and SVM artifact that is assigned to it. The object's compound (final) risk level is then set to the highest value of its assigned risk levels.

Retention Duration for S3 Metadata

Macie stores metadata about your S3 objects for the default duration of 1 month. You can extend this duration up to 12 months.

Protecting Data with Amazon Macie

Topics

Macie can help you monitor how your sensitive and business-critical data stored in the cloud is being used. Macie applies artificial intelligence to understand historical data access patterns and automatically assesses activity of users, applications, and service accounts. This can help you detect unauthorized access and avoid data leaks.

Once you enable Macie, it uses the following automated methods to protect your data.

AWS CloudTrail Events

Macie analyzes and processes a subset of data that CloudTrail logs and management events (API calls) that can occur in your infrastructure. Macie designates a risk level between 1 and 10 for each of the supported CloudTrail events.

You can't modify existing or add new CloudTrail events to the list that Macie manages. You can enable or disable the supported CloudTrail events, thus instructing Macie to either include or exclude them in its data security process.

To view, enable, or disable supported CloudTrail events

1. In the Macie console, navigate to the **Settings** page.
2. In the **Protect data** section, choose **AWS CloudTrail events**.
3. Choose any of the listed events to view its details.

   To enable or disable an event, on its details page, use the **Enabled/Disabled** dropdown and then choose **Save**.

AWS CloudTrail Errors

Macie analyzes and processes errors that can occur when a subset of data that CloudTrail logs and management events (API calls) take place in your infrastructure. Macie designates a risk level between 1 and 10 for each of the supported CloudTrail errors.

You can't modify existing or add new CloudTrail errors to the list that Macie manages. You can enable or disable the supported CloudTrail errors, thus instructing Macie to either include or exclude them in its data security process.

To view, enable, or disable supported CloudTrail errors

1. In the Macie console, navigate to the **Settings** page.
2. In the **Protect data** section, choose **AWS CloudTrail errors**.
3. Choose any of the listed errors to view its details.

To enable or disable an error, on its details page, use the **Enabled/Disabled** dropdown and then choose **Save**.

Using the Amazon Macie Dashboard

The Macie **Dashboard** draws a comprehensive picture of all of your data and activity that Macie monitors. This topic describes the metrics and views that you can use in the **Dashboard** to view your monitored data grouped by various interest points. Each metric and view provides you with one or more ways of navigating to the Macie console's **Research** tab. There you can construct and run queries in the query parser and conduct in-depth investigative research of your data and activity that Macie monitors.

Dashboard Metrics

The following **Dashboard** metrics enable you to view your monitored data grouped by several key interest points:

- **High-risk S3 objects** – While classifying data (p. 14), Macie assigns a risk value to each monitored data object. This is Macie's way of helping you identify and prioritize your sensitive data over other, less business-critical data. This metric allows you to see all of your Macie-monitored data objects with a risk levels of 8 through 10.
- **Total event occurrences** – As part of securing data (p. 32), Macie analyzes and processes events (API calls) logged by AWS CloudTrail that occur within your infrastructure. This metric provides the total count of all of the event occurrences monitored by Macie that took place within your infrastructure since you enabled Macie.
- **Total user sessions** – A user session is a 5-minute aggregate of CloudTrail data. This metric provides the total count of all user sessions of CloudTrail data that Macie analyzed and processed since it was enabled.

Dashboard Views

Follow this procedure to use the predefined Macie **Dashboard** views and generate distinct subsets of your data and activity monitored by Macie.

To use Macie dashboard views

1. Choose the corresponding icon to select any of the following views to display various subsets of your data and activity monitored by Macie:
   - S3 objects for a selected time range (p. 35)
   - S3 objects (p. 35)
   - S3 objects by PII (p. 36)
   - S3 public objects by buckets (p. 36)
   - S3 objects by ACL (p. 37)
   - CloudTrail events and associated users (p. 37)
   - CloudTrail errors and associated users (p. 38)
   - Activity location (p. 39)
   - AWS CloudTrail events (p. 39)
   - Activity ISPs (p. 40)
   - AWS CloudTrail user identity types (p. 40)
2. If present in the selected view, locate and move the **Minimum risk** slider to the desired value. The **Minimum risk** slider enables you to view only items with the assigned risk equal to and greater than the selected value.

S3 Objects for Selected Time Range

This view provides a visual representation of your monitored S3 objects that match the following search criteria:

- At least one of the object's assigned themes is of the top 20 most frequently assigned themes
- The object's assigned risk is equal to or greater than the value selected on the **Minimum risk** slider
- The object was last modified during one of the following time ranges:
  - The past 6 months
  - Between the date when Macie was enabled and a date six months before today

To navigate from this view to the **Research** tab, select (double-click) any of the squares that represent the displayed time ranges or themes. Your selection is automatically translated into a query that appears in the query parser in the **Research** tab.

You can follow this sample procedure.

1. In the Macie **Dashboard**, select the **S3 objects over selected time range** view.
2. Set the **Minimum risk** slider to 5.
3. In the generated graph, double-click the square next to **Range: 0 - 6 months ago**.

   As a result, you're redirected to the **Research** tab with the following query, which automatically appears in the query parser:

   ```
   themes:* AND dlp_risk:[5 TO *] AND @timestamp:[now-6M/M TO now]
   ```

   This query matches your selection to view the S3 objects monitored by Macie that are assigned one or more of the top 20 most frequently assigned themes, that have an assigned risk of 5 or higher, and that were last modified at some point in the past 6 months. The results of this query also appear. You can modify the query result controls available on the **Research** tab, run the query again, and conduct in-depth investigative research of the generated results. For more information, see Researching Through Data Monitored by Amazon Macie (p. 53).

S3 Objects

This view provides the complete list of your S3 objects monitored by Macie, grouped by the assigned themes. For each theme, a percentage that this theme represents of the total number of your S3 objects monitored by Macie is displayed, as well as the total count of the S3 objects that were assigned this theme.

To navigate from this view to the **Research** tab, choose the looking glass icon next to any of the displayed themes. Your selection is automatically translated into a query that appears in the query parser in the **Research** tab.

You can follow this sample procedure.

1. In the Macie **Dashboard**, select the **S3 objects** view.
2. From the generated list of S3 objects, choose the looking glass icon next to, for example, **json/aws_cloudtrail_logs**.

   As a result, you're redirected to the **Research** tab with the following query, which automatically appears in the query parser:

   ```
   themes:"json/aws_cloudtrail_logs"
   ```

This query matches your selection to view the S3 objects monitored by Macie with the assigned theme of **json/aws_cloudtrail_logs**. The results of this query also appear. You can modify the query result controls available on the **Research** tab, run the query again, and conduct in-depth investigative research of the generated results. For more information, see Researching Through Data Monitored by Amazon Macie (p. 53).

S3 Objects by PII

This view provides the following lists:

- **S3 objects by PII priority**

  This is a complete list of your S3 objects that contain PII artifacts, grouped by the PII priority assigned by Macie. For each PII priority level, a percentage that the number of objects with this level represents of the total number of the S3 objects with PII artifacts is displayed, as well as the total count of the S3 objects with this PII priority level.
- **S3 objects by PII types**

  This is a complete list of your S3 objects that contain PII artifacts, grouped by the PII artifact types. For each PII artifact type, a percentage that the number of objects with PII artifacts of this type represents of the total number of the S3 objects with PII artifacts is displayed, as well as the total count of the S3 objects with PII artifacts of this type.

For more information about PII-based object classification, see Classifying Data with Amazon Macie (p. 14).

To navigate from this view to the **Research** tab, choose the looking glass icon next to any of the displayed PII impacts or PII types. Your selection is automatically translated into a query that appears in the query parser in the **Research** tab.

You can follow this sample procedure.

1. In the Macie **Dashboard**, select the **S3 objects by PII** view.
2. For example, let's generate a list of S3 objects with low PII priority. In the **S3 objects by PII priority** list, choose the looking glass icon next to the low PII priority.

   As a result, you're redirected to the **Research** tab with the following query, which automatically appears in the query parser:

   ```
   pii_impact:"low"
   ```

   The results of this query also appear. You can modify the query result controls available on the **Research** tab, run the query again, and conduct in-depth investigative research of the generated results. For more information, see Researching Through Data Monitored by Amazon Macie (p. 53).

S3 Public Objects by Buckets

This is a complete list of your public S3 objects grouped by the buckets that they're stored in. For each bucket, a percentage that this bucket's objects represent of the total number of your S3 objects managed by Macie is displayed, as well as the total count of the S3 objects that are stored in this bucket.

To navigate from this view to the **Research** tab, choose the looking glass icon next to any of the displayed buckets. Your selection is automatically translated into a query that appears in the query parser in the **Research** tab.

S3 Objects by ACL

This view provides the following lists:

- **S3 objects by ACL URIs**

  This is a complete list of URIs that appear in access control lists (ACLs) that are attached to your S3 objects. For each URI, a percentage that the number of objects with ACLs attached that contain this URI represents of the total number of the S3 objects with ACLs attached is displayed, as well as the total count of the S3 objects with ACLs attached that contain this URI.
- **S3 objects by ACL display names**

  This is a complete list of user display names that appear in ACLs that are attached to your S3 objects. For each display name, a percentage that the number of objects with ACLs attached that contain this display name represents of the total number of the S3 objects with ACLs attached is displayed, as well as the total count of the S3 objects with ACLs attached that contain this display name.
- **S3 objects by ACL permissions**

  This is a complete list of access permissions that appear in ACLs that are attached to your S3 objects. For each permissions level, a percentage that the number of objects with ACLs attached that contain this permission level represents of the total number of the S3 objects with ACLs attached is displayed, as well as the total count of the S3 objects with ACLs attached that contain this permission level.

To navigate from this view to the **Research** tab, choose the looking glass icon next to any of the displayed URIs, ACL display names, and ACL permissions. Your selection is automatically translated into a query that is displayed in the query parser in the **Research** tab.

You can follow this sample procedure.

1. In the Macie **Dashboard**, select the **S3 objects by ACL** view.
2. For example, let's generate a list of S3 objects with attached ACLs that contain full control permissions. In the **S3 objects by ACL permissions** list, choose the looking glass icon next to the **FULL_CONTROL** permission.

   As a result, you're redirected to the **Research** tab with the following query, which automatically appears in the query parser.

   ```
   object_acl.Grants.Permission:"FULL_CONTROL"
   ```

   The results of this query also appear. You can modify the query result controls available on the **Research** tab, run the query again, and conduct in-depth investigative research of the generated results. For more information, see Researching Through Data Monitored by Amazon Macie (p. 53).

CloudTrail Events and Associated Users

This view provides the following lists:

- **AWS CloudTrail events**

  This is a visual representation of the top 20 (by assigned risk and based on the value selected on the **Minimum risk** slider) CloudTrail data and management events that occurred during the last 60 days. You can use the available **Daily** or **Weekly** radio buttons to modify the graph to view daily or weekly results.

  To navigate from this view to the **Research** tab, select (double-click) any square that represents a particular event that you want to investigate further. The number in parentheses next to the event

name represents the number of user sessions (5-minute aggregates of CloudTrail data) that this event is present in. In the **Research** tab, your selection is automatically translated into a query that appears in the query parser.

- **AWS CloudTrail associated users**

  This is a visual representation of the users associated with the top 20 (by assigned risk and based on the value selected on the **Minimum risk** slider) CloudTrail data and management events that occurred during the last 60 days. You can use the available **Daily** or **Weekly** radio buttons to modify the graph to view daily or weekly results.

  To navigate from this view to the **Research** tab, select (double-click) any square that represents a particular error that you want to investigate further. The number in parentheses next to the user name represents the number of user sessions (5-minute aggregates of CloudTrail data) that this user is associated with. In the **Research** tab, your selection is automatically translated into a query that appears in the query parser.

You can follow this sample procedure.

1. In the Macie **Dashboard**, select the **CloudTrail events and associated users** view.
2. Set the **Minimum risk** slider to 1.
3. For example, let's generate a list of user sessions that the **PutRestApi** event is present in. Double-click the square next to **PutRestApi**.

   As a result, you're redirected to the **Research** tab with the following query, which automatically appears in the query parser.

   ```
   eventNameIsp.key.keyword:"PutRestApi" AND @timestamp:[now-60d TO now]
   ```

   The results of this query also appear. You can modify the query result controls available on the **Research** tab, run the query again, and conduct in-depth investigative research of the generated results. For more information, see Researching Through Data Monitored by Amazon Macie (p. 53).

CloudTrail Errors and Associated Users

This view provides the following lists:

- **AWS CloudTrail errors**

  This is a visual representation of the top 20 (by assigned risk and based on the value selected on the **Minimum risk** slider) CloudTrail errors that occurred during the last 60 days. You can use the available **Daily** or **Weekly** radio buttons to modify the graph to view daily or weekly results.

  To navigate from this view to the **Research** tab, select (double-click) any square that represents a particular error that you want to investigate further. The number in parentheses next to the error name represents the number of user sessions (5-minute aggregates of CloudTrail data) that this error is present in. In the **Research** tab, your selection is automatically translated into a query that appears in the query parser.

- **AWS CloudTrail associated users**

  This is a visual representation of the users associated with the top 20 (by assigned risk and based on the value selected on the **Minimum risk** slider) CloudTrail errors that occurred during the last 60 days. You can use the available **Daily** or **Weekly** radio buttons to modify the graph to view daily or weekly results.

  To navigate from this view to the **Research** tab, select (double-click) any square that represents a particular error that you want to investigate further. The number in parentheses next to the user

name represents the number of user sessions (5-minute aggregates of CloudTrail data) that this user is associated in. In the **Research** tab, your selection is automatically translated into a query that appears in the query parser.

You can follow this sample procedure.

1. In the Macie **Dashboard**, select the **CloudTrail errors and associated users** view.
2. Set the **Minimum risk** slider to 1.
3. For example, let's generate a list of user sessions that the **Client.InvalidPermission.NotFound** error is present in. Double-click the square next to **Client.InvalidPermission.NotFound**.

   As a result, you're redirected to the **Research** tab with the following query, which automatically appears in the query parser.

   ```
   eventNameErrorCode.secondary:"Client.InvalidPermission.NotFound" AND
   @timestamp:[now-60d TO now]
   ```

   The results of this query also appear. You can modify the query result controls available on the **Research** tab, run the query again, and conduct in-depth investigative research of the generated results. For more information, see Researching Through Data Monitored by Amazon Macie (p. 53).

Activity location

This view includes a map that shows the locations of activity that Macie is monitoring for a selected time period. To view details, use the available time period pull-down menu (past 15 days, past 30 days, past 90 days, or past year) and then choose any location pin.

To navigate from this view to the **Research** tab, choose the number of events that appears in a tool tip window for a location pin. In the **Research** tab, your selection is automatically translated into a query that appears in the query parser. For example, you can autogenerate the following query to display a list of user sessions that occurred in the past 15 days in Seattle.

```
geoLocation.key:"Seattle:UnitedStates:47.6145:-122.348" AND @timestamp:[now-15d
TO now]
```

You can modify the query result controls available on the **Research** tab, run the query again, and conduct in-depth investigative research of the generated results. For more information, see Researching Through Data Monitored by Amazon Macie (p. 53).

AWS CloudTrail Events

AWS CloudTrail events

This view provides the complete list of your CloudTrail data and management events monitored by Macie. For each event, the total count of the user sessions (5-minute integrations of CloudTrail data) that this event is present in and the percentage that this total represents of the total number of user sessions appears.

To navigate from this view to the **Research** tab, choose the looking glass icon next to any of the displayed events. Your selection is automatically translated into a query that appears in the query parser in the **Research** tab. For example, you can autogenerate the following query to view all user sessions that the **AssumeRole** event is present in.

```
eventNameIsp.key.keyword:"AssumeRole"
```

You can modify the query result controls available on the **Research** tab, run the query again, and conduct in-depth investigative research of the generated results. For more information, see Researching Through Data Monitored by Amazon Macie (p. 53).

Activity ISPs

Activity ISPs

This view provides the complete list of your CloudTrail data and management events monitored by Macie, grouped by the associated internet service providers (ISPs). For each ISP, the total count of the user sessions (5-minute integrations of CloudTrail data) that this ISP is present in and the percentage that this total represents of the total number of user sessions appears.

To navigate from this view to the **Research** tab, choose the looking glass icon next to any of the displayed themes. Your selection is automatically translated into a query that appears in the query parser in the **Research** tab. For example, you can autogenerate the following query to view all user sessions that are associated with Amazon.

```
eventNameIsp.secondary.keyword:"Amazon"
```

You can modify the query result controls available on the **Research** tab, run the query again, and conduct in-depth investigative research of the generated results. For more information, see Researching Through Data Monitored by Amazon Macie (p. 53).

AWS CloudTrail User Identity Types

This view provides the complete list of your CloudTrail data and management events monitored by Macie, grouped by the user identity type (for more information, see the definition for *user* in Concepts and Terminology (p. 3)). For each user identity type, the total count of the user sessions (5-minute integrations of CloudTrail data) that this user identity type is present in and the percentage that this total represents of the total number of user sessions is displayed.

To navigate from this view to the **Research** tab, choose the looking glass icon next to any of the displayed themes. Your selection is automatically translated into a query that appears in the query parser in the **Research** tab. For example, you can autogenerate the following query to view all user sessions that contain requests that were originated by the **AssumedRole** user identity type.

```
userIdentityType.key:"AssumedRole"
```

You can modify the query result controls available on the **Research** tab, run the query again, and conduct in-depth investigative research of the generated results. For more information, see Researching Through Data Monitored by Amazon Macie (p. 53).

Amazon Macie Alerts

An *alert* is a notification about a potential security issue discovered by Amazon Macie. This section describes the following information:

Topics

Basic and Predictive Macie Alerts

Macie generates two types of alerts:

- **Basic alerts** – Alerts generated by the security checks that Macie performs. There are two types of basic alerts in Macie:
  - Managed (curated by Macie) basic alerts that you can't modify. You can enable or disable the existing managed basic alerts.

    Note
    You can identify managed basic alerts by the value of `MacieDefault` in the **Created by** field in the **Basic alerts** list in the **Settings** tab.
  - Custom basic alerts that you can create and modify to your exact specifications. For more information, see Adding New and Editing Existing Custom Basic Alerts (p. 44).
- **Predictive alerts** – Automatic alerts based on activity in your AWS infrastructure that deviates from the established normal activity baseline. More specifically, Macie continuously monitors activity in your AWS infrastructure and builds a model of the normal behavior. Then it looks for deviations from that normal baseline, and when it detects such activity, it generates automatic predictive alerts. For example, a user uploading or downloading a large number of S3 objects in a day might trigger an alert if that user typically downloads one or two S3 objects in a week.

Alert Categories in Macie

Macie's basic alerts (managed and custom) can be of the following categories:

- **Configuration compliance** – Related to compliance-controlled content, policy, configuration settings, control and data plane logging, and patch level.
- **Data compliance** – Related to the discovery of compliance or security-controlled content, such as the existence of Personally Identifiable Information (PII), or access credentials.
- **File hosting** – Related to you hosting possible malware, unsafe software, or attackers' command and control infrastructure through compromised hosts or storage services.

- **Service disruption** – Configuration changes that can lead to you being unable to access resources in your own environment.
- **Ransomware** – Potentially malicious software or activity designed to block your access to your own computer system until a sum of money is paid.
- **Suspicious access** – Access to your resources from a risky anomalous IP address, user, or system, such as an attacker masquerading their connection through a compromised host.
- **Identity enumeration** – A series of API calls or accesses enumerating access levels to your systems that can possibly indicate the early stages of an attack or compromised credentials.
- **Privilege escalation** – Successful or unsuccessful attempts to gain elevated access to resources that are normally protected from an application or user, or attempts to gain access to your system or network for an extended period of time.
- **Anonymous access** – Attempted access to your resources from an IP address, user, or service with the intent to hide a user's true identity. Examples include the use of proxy servers, virtual private networks, and other anonymity services such as Tor.
- **Open permissions** – Identification of sensitive resources protected by potentially overly permissive (and thus risky) access control mechanisms.
- **Location anomaly** – An anomalous and risky location of the access attempt to your sensitive data.
- **Information loss** – An anomalous and risky access to your sensitive data.
- **Credentials loss** – Possible compromise of your credentials.

To view a list of your existing alerts of a particular category, choose that category from the **Categories** list on the Macie console's **Alerts** tab.

Severity Levels for Alerts in Macie

Each Macie alert has an assigned severity level. This reduces the need to prioritize one alert over another in your analyses. It can also help you determine your response when an alert highlights a potential problem. **Critical**, **High**, **Medium**, and **Low** levels indicate a security issue that can result in compromised information confidentiality, integrity, and availability in your infrastructure. The **Informational** level highlights a security configuration detail of your infrastructure that Macie monitors. The following are recommended ways to respond to each level:

- **Critical** – Describes a security issue that can result in a compromise of the information confidentiality, integrity, and availability in your infrastructure. We recommend that you treat this security issue as an emergency and implement an immediate remediation. The main difference between a **Critical** and **High** severity is that a **Critical** severity alert might be informing you of a security compromise of a large number of your resources or systems. A **High** severity alert is informing you of a security compromise of one or several of your resources or systems.
- **High** – Describes a security issue that can result in a compromise of the information confidentiality, integrity, and availability in your infrastructure. We recommend that you treat this security issue as an emergency and implement an immediate remediation.
- **Medium** – Describes a security issue that can result in a compromise of the information confidentiality, integrity, and availability in your infrastructure. We recommend that you fix this issue at the next possible opportunity, for example, during your next service update.
- **Low** – Describes a security issue that can result in a compromise of the information confidentiality, integrity, and availability in your infrastructure. We recommend that you fix this issue as part of one of your future service updates.
- **Informational** – Describes a particular security configuration detail of your infrastructure. Based on your business and organization goals, you can either note this information or use it to improve the security of your systems and resources.

Locating and Analyzing Macie Alerts

You can use the following procedure to locate and analyze existing alerts.

1. To view your generated alerts (including **Active** and **Archived** basic or predictive alerts), in the Macie console, navigate to the **Alerts** page.

   Each alert has a summary section that contains the following information:

   - Alert severity, which can be **Critical**, **High**, **Medium**, **Low**, or **Informational**. For more information, see Severity Levels for Alerts in Macie (p. 42).
   - A timestamp that indicates when the alert was generated or last updated.
   - The alert category. For more information, see Alert Categories in Macie (p. 41).
   - One of the following:
     - If the alert's index is **CloudTrail data**, a user that engaged in the activity that prompted Macie to generate the alert. For more information, see the definition of *user* in the context of Macie in Concepts and Terminology (p. 3).
     - If the alert's index is **S3 bucket properties** or **S3 objects**, a bucket name that was involved in or that contains the objects that were involved in the activity that prompted Macie to generate the alert.

       Important
       In Macie, each alert is based on one of the following:

       - For the alerts with the index of **CloudTrail data**, only one user: the IAM identity whose activity prompted Macie to generate the alert.
       - For the alerts with the index of **S3 bucket properties** or **S3 objects**, only one S3 bucket that was involved in or that contains objects that were involved in the activity that prompted Macie to generate the alert.
   - The number of comments that were left on the alert.
   - The total number of results, which can consist of a list of user sessions, or a list of S3 buckets, or a list of S3 objects that match the query that is included in the definition of the alert. For more information, see Adding New and Editing Existing Custom Basic Alerts (p. 44).
   - The number of views on the alert.
   - The AWS Region where the activity captured in this alert took place.
2. To analyze any alert further, choose the alert to expand its details pane. The following information is included in the alert details:

   - The alert summary that includes the description and the total number of results: a number of user sessions, S3 buckets, or S3 objects that match the query that is included in the definition of the alert.
   - A list of the alert results. This is a list of user sessions, S3 buckets, or S3 objects, depending on the index that is specified in the definition for this alert. For more information, see Adding New and Editing Existing Custom Basic Alerts (p. 44).
     - If you specified **CloudTrail data** as the index, the alert details contain a list of user sessions that match the query specified in the alert definition for a particular user.
     - If you specified **S3 buckets** as the index, the alert details contain a list of S3 buckets that match the query specified in the alert definition for a particular user.
     - If you specified **S3 objects** as the index, the alert details contain a list of S3 objects that match the query specified in the alert definition for a particular user.

   You can choose each result to examine it and view all its fields. For more information, see the Researching AWS Data, Researching S3 Bucket Properties Data, or Researching S3 Objects Data sections in Researching Through Data Monitored by Amazon Macie (p. 53).

You can also use the **Research** looking glass icon to navigate to the **Research** tab and view the results of a particular alert there. The query parser in the **Research** tab is then prepopulated with the query that can be used to generate these results.

Adding New and Editing Existing Custom Basic Alerts

You can use the following procedure to add new and edit existing custom basic alerts.

1. In the Macie console, navigate to the **Settings** page and choose the icon for **Basic alerts**.
2. On the **Basic alerts** page, either choose the edit icon for the alert that you want to modify or, to add a basic alert, choose **Add new**.
3. Do one of the following:

   - If you're editing the existing alert, make your changes, including enabling or disabling the alert, and then choose **Save**.
   - If you're adding a new alert, on the **Basic alert definition** page, specify the following:
     - Alert title – For example, "An S3 bucket has an S3 bucket policy or S3 ACL that grants read rights to everyone."
     - Description for the alert – For example, "An S3 bucket policy or S3 ACL on an S3 bucket contains a clause that effectively grants read access to any user. We recommend that you audit this S3 bucket and its data and confirm that this is intentional.
     - Alert category – For more information, see Alert Categories in Macie (p. 41).
     - Alert query – A query that describes the activity that you want Macie to generate an alert about. For example, `s3_world_readability:"true"`. This query looks for an S3 bucket policy or S3 ACL policy on an S3 bucket that grants read access to any user. For more information about constructing queries, see Constructing Queries in Macie (p. 53).

       Note
       You can use the looking glass icon next to an existing alert to navigate to the **Research** tab. This alert's query automatically appears in the **Query Parser**, and the results of this query appears in the **Research** tab.
     - Query index – The repository of data against which Macie will run the query specified in this alert. You can select either CloudTrail data, S3 buckets, or S3 objects. Depending on your selection, the alert will contain a list of CloudTrail user sessions (5-minute aggregates of raw CloudTrail data), S3 buckets, or S3 objects that match the activity that your alert defines.
     - A minimum number of activity matches that must occur before an alert is generated.
     - Alert severity – For more information, see Severity Levels for Alerts in Macie (p. 42)
     - Whitelisted users or whitelisted buckets, depending on the selected alert index. If you whitelist a user or a bucket, Macie doesn't generate an alert for this user or bucket when they're involved in the activity that the alert defines.

       Important
       In Macie, each alert is based on one of the following:
       - For the alerts with the index of **CloudTrail data**, only one user: the IAM identity whose activity prompted Macie to generate the alert.
       - For the alerts with the index of **S3 bucket properties** or **S3 objects**, only one S3 bucket that was involved in or that contains objects that were involved in the activity that prompted Macie to generate the alert.

       When whitelisting a user in a basic alert with the index of CloudTrail data, you must use a special Macie format called `macieUniqueId`. Examples include `123456789012:root`,

`123456789012:user/Bob`, and `123456789012:assumed-role/Accounting-Role/Mary`, depending on the identity type of the user that you want to whitelist. For more information, see the definition of *user* in Analyzing Amazon Macie–Monitored Data by User Activity (p. 48).

- Specify whether this alert is enabled or disabled.

Working with Existing Alerts

You can use the following procedure to archive or unarchive alerts or to choose edit the existing basic alerts.

1. In the Macie console, navigate to the **Alerts** page and locate the alert that you want to archive, unarchive (if it's an archived alert), or edit.
2. Choose the down arrow in the alert summary pane and then choose either of the following:

   - **Archive**

     Note
     Or **Unarchive** if this is an archived alert.
   - **Edit basic alert**

     Important
     This option isn't available for predictive alerts. You can't edit predictive alerts, which Macie automatically generates based on activity in your AWS infrastructure that deviates from the established normal activity baseline. For more information, see Basic and Predictive Macie Alerts (p. 41).

Group Archiving Alerts

You can use the following procedure to group archive alerts.

1. In the Macie console's **Alerts** page, choose **Group Archive**.
2. In the **Group archive** window, use the available settings to archive or unarchive multiple alerts at the same time.

Whitelisting Users or Buckets for Basic Alerts

Macie allows you to whitelist users (if the alert's index is **CloudTrail data**) and buckets (if the alert's index is **S3 bucket properties** or **S3 objects**) for both alerts managed by Macie and custom basic alerts.

Note
Macie doesn't allow you to whitelist users or buckets for predictive alerts.

You can use the following procedure to whitelist a specific user or a specific bucket that engaged in or was involved in the activity that prompted Macie to generate a specific alert.

Important
In Macie, each alert is based on one the following:

- For the alerts with the index of **CloudTrail data**, only one user: the IAM identity whose activity prompted Macie to generate the alert.
- For the alerts with the index of **S3 bucket properties** or **S3 objects**, only one S3 bucket that was involved in or that contains objects that were involved in the activity that prompted Macie to generate the alert.

To whitelist users or S3 buckets for custom basic alerts using the Alerts tab

1. In the Macie console's **Alerts** tab, locate the custom basic alert for which you want to whitelist a user or S3 bucket listed in the alert's summary.
2. Choose the down arrow in the alert summary pane and then choose **Whitelist user** (if this alert's index is **CloudTrail data**) or **Whitelist bucket** (if the alert's index is **S3 bucket properties** or **S3 objects**).
3. In the **Whitelist user** (or **Whitelist bucket**) window, verify the user or bucket that you want to whitelist (automatically preselected and matching the user or bucket listed in the alert's summary) and then choose **Submit**.

You can use the following procedure to whitelist multiple users or buckets at the same time for custom basic alerts.

To whitelist users or S3 buckets for custom basic alerts using the Settings tab

1. In the Macie console's **Settings** tab, choose **Basic alerts** and then locate the custom basic alert for which you want to whitelist users or S3 buckets.
2. Choose the edit icon next to the alert.
3. Specify users or S3 buckets that you want to whitelist in either **Whitelisted users** (if this alert's index is **CloudTrail data**) or **Whitelisted buckets** (if the alert's index is **S3 bucket properties** or **S3 objects**) fields and choose **Save**.

   Note
   When whitelisting a user in a basic alert with the index of CloudTrail data, you must use a special Macie format called `macieUniqueId`: Examples include `123456789012:root`, `123456789012:user/Bob`, and CloudTrail, depending on the identity type of the user you want to whitelist. For more information, see the definition of the user concept in Analyzing Amazon Macie–Monitored Data by User Activity (p. 48).

Whitelist users or S3 buckets for Macie-managed basic alerts

1. In the Macie console's **Alerts** tab, locate the basic alert managed by Macie that you want to whitelist users or S3 buckets for.
2. Choose the down arrow in the alert summary pane and then choose **Whitelist user** (if this alert's index is **CloudTrail data**) or **Whitelist bucket** (if the alert's index is **S3 bucket properties** or **S3 objects**).
3. In the **Whitelist user** or **Whitelist bucket** window, select the **Clone and disable the default managed alert** check box and choose **Submit**.
4. Navigate to the Macie console's **Settings** tab.

   The original managed alert that you worked with in the previous step is now disabled. This alert has also been cloned into a new custom basic alert. For example, if your original managed basic alert was called "An S3 bucket has an S3 bucket policy or S3 ACL that grants read rights to everyone," this alert is now disabled, and a custom basic alert called "An S3 bucket has an S3 bucket policy or S3 ACL that grants read rights to everyone (modified)" is created (cloned).
5. Choose the edit icon next to the cloned custom basic alert.
6. Specify users or S3 buckets that you want to whitelist in either **Whitelisted users** (if this alert's index is **CloudTrail data**) or **Whitelisted buckets** (if the alert's index is **S3 bucket properties** or **S3 objects**) fields and choose **Save**.

   Note
   When whitelisting a user in a basic alert with the index of CloudTrail data, you must use a special Macie format called `macieUniqueId`: Examples include `123456789012:root`, `123456789012:user/Bob`, and `123456789012:assumed-role/Accounting-Role/`

`Mary`, depending on the identity type of the user that you want to whitelist. For more information, see the definition of *user* in Analyzing Amazon Macie–Monitored Data by User Activity (p. 48).

Analyzing Amazon Macie–Monitored Data by User Activity

The **Users** tab can help you draw a comprehensive picture of all of the data and activity monitored by Macie for a particular selected user. This topic describes how to search for the users whose activity you want to investigate further in the **Users** tab. It also describes the views that you can use in this tab to see the selected users' monitored data grouped by various interest points. Each view provides you with one or more ways of navigating to the Macie console's **Research** tab. There you can construct and run queries in the query parser and conduct in-depth investigative research of the data and activity monitored by Macie for the selected users.

Topics

MacieUniqueID

In the context of Macie, a user is the AWS Identity and Access Management (IAM) identity that makes a particular request. Macie uses the AWS CloudTrail `userIdentity` element to distinguish the following user types. For more information, see CloudTrail userIdentity Element.

- Root – The request was made with your AWS account credentials.
- IAM user – The request was made with the credentials of an IAM user.
- Assumed role – The request was made with temporary security credentials that were obtained with a role via a call to the AWS Security Token Service (AWS STS) `AssumeRole` API operation.
- Federated user – The request was made with temporary security credentials that were obtained via a call to the AWS STS `GetFederationToken` API operation.
- AWS account – The request was made by another account.
- AWS service – The request was made by an account that belongs to an AWS service.

When specifying a user in the Macie console, you must use a special Macie format called `macieUniqueId`. Examples of specifying a user include searching for a user in the **Users** tab, constructing a query in the **Research** tab, and whitelisting a user in a basic alert with the index of **CloudTrail data**. The `macieUniqueId` is a combination of the IAM `UserIdentity` element and the `recipientAccountId`. For more information, see CloudTrail userIdentity Element and the definition of `recipientAccountId` in CloudTrail Record Contents.

The following examples list various structures of `macieUniqueId`, depending on the user identity type.

userIdentity	MacieUniqueId
`"userIdentity": { "type": "AssumedRole" "arn": "arn:aws:sts::123456789012:assumed-role/Accounting-Role/Mary" }`	`123456789012:assumed-role/accounting-role`

userIdentity	MacieUniqueId

<pre>"userIdentity": {
    "type": "IAMUser",
    "arn":
 "arn:aws:iam::123456789012:user/
Bob",
    "userName": "Bob"
}</pre>
	123456789012:user:bob

<pre>"userIdentity": {
    "type": "FederatedUser"
    "arn":
 "arn:aws:sts::123456789012:federated-
user/Alice",
    "principalId":
 "123456789012:Alice",
}</pre>
	123456789012:federated-user:alice

<pre>"recipientAccountId":
 "123456789012",
"userIdentity": {
    "type": "AWSAccount"
    "accountId":
 "ANONYMOUS_PRINCIPAL",
}</pre>
	123456789012:ANONYMOUS_PRINCIPAL

<pre>"macieUniqueId":
 "123456789012:root:root",
"userIdentity": {
    "type": "Root"
    "sourceARN":
 "arn:aws:iam::123456789012:root",
}</pre>
	123456789012:root:root

<pre>"userIdentity": {
    "invokedBy":
 "codepipeline.amazonaws.com",
    "type": "AWSService"
}
"recipientAccountId":
 "123456789012",</pre>
	123456789012:codepipeline.amazonaws.com

<pre>"recipientAccoundId":
 "123456789012",
"userIdentity": {
    "type": "AWSAccount"
    "accountId":
 "987654321098",
    "principalId":
 "AIDABCDEFGHI123456XYZ",
}</pre>
	123456789012:AIDABCDEFGHI123456XYZ

User Categories in Macie

Based on their activity (API calls), users in Macie are grouped into the following categories:

- **Platinum** – These IAM users or roles have a history of making high-risk API calls indicative of an administrator or root user, such as creating users, authorizing security group ingress, or updating policies. These accounts should be monitored closely for signs of account compromise.
- **Gold** – These IAM users or roles have a history of making infrastructure-related API calls indicative of a power user, such as running instances or writing data to Amazon Simple Storage Service (Amazon S3). These accounts should be monitored closely for signs of account compromise.
- **Silver** – These IAM users or roles have a history of issuing high quantities of medium-risk API calls, such as `Describe*` and `List*` operations, or read-only access requests to Amazon S3.
- **Bronze** – These IAM users or roles typically execute lower quantities of `Describe*` and `List*` API calls in the AWS environment.

Investigating Users

Follow this procedure to generate a comprehensive picture of all of the data and activity monitored by Macie for the specified user.

1. In the Macie console's **Users** tab, specify a user name in the **Search** field and press Enter.

   Note
   When specifying a user, you must use a special Macie format called **macieUniqueId**: for example, `123456789012:root`, `123456789012:user/Bob`, or `123456789012:assumed-role/Accounting-Role/Mary`, depending on the identity type of the user that you want to whitelist. For more information, see the definition of *user* in Concepts and Terminology (p. 3).
2. When the user data is generated, choose the corresponding icon to select any of the following views to display various subsets of this user's data and activity that Macie monitors:

   - High-risk CloudTrail events (p. 50)
   - High-risk CloudTrail errors (p. 51)
   - Activity location (p. 51)
   - CloudTrail events (p. 51)
   - Activity ISPs (p. 51)
   - CloudTrail user identity types (p. 51)
3. If present in the selected view, locate and move the **Minimum risk** slider to the desired value. The **Minimum risk** slider enables you to view only items with the assigned risk equal to and greater than the selected value.

High-Risk CloudTrail Events

This view provides a visual representation of the top 20 (by assigned risk and based on the value selected on the **Minimum risk** slider) CloudTrail data and management events that occurred during the last 60 days for the selected user. Use the available **Daily** or **Weekly** radio buttons to modify the graph to view daily or weekly results.

To navigate from this view to the **Research** tab, select (double-click) any square that represents a particular event that you want to investigate further. The number in parentheses next to the event name represents the number of user sessions (5-minute aggregates of CloudTrail data) that this event is present in. In the **Research** tab, your selection is automatically translated into a query that appears

in the query parser. For more information, see Researching Through Data Monitored by Amazon Macie (p. 53).

High-Risk CloudTrail Errors

This view provides a visual representation of the top 20 (by assigned risk and based on the value selected on the **Minimum risk** slider) CloudTrail errors that occurred during the last 60 days for the selected user. You can use the available **Daily** or **Weekly** radio buttons to modify the graph to view daily or weekly results.

To navigate from this view to the **Research** tab, select (double-click) any square that represents a particular error that you would like to investigate further. The number in parentheses next to the error name represents the number of user sessions (5-minute aggregates of CloudTrail data) that this error is present in. In the **Research** tab, your selection is automatically translated into a query that appears in the query parser. For more information, see Researching Through Data Monitored by Amazon Macie (p. 53).

Activity Location

This view includes a map that shows the locations of activity that Macie is monitoring for a selected time period for the specified user. To view details, use the available time period dropdown (past 15 days, past 30 days, past 90 days, or past year) and then choose any location pin.

To navigate from this view to the **Research** tab, choose the number of events that appears in a tool tip window for a location pin. In the **Research** tab, your selection is automatically translated into a query that appears in the query parser. For more information, see Researching Through Data Monitored by Amazon Macie (p. 53).

CloudTrail Events

This view provides the complete list of CloudTrail data and management events monitored by Macie for the specified user. For each event, the total count of the user sessions (5-minute integrations of CloudTrail data) that this event is present in, and the percentage that this total represents of the total number of user sessions is displayed.

To navigate from this view to the **Research** tab, choose the looking glass icon next to any of the displayed events. Your selection is automatically translated into a query that appears in the query parser in the **Research** tab. For more information, see Researching Through Data Monitored by Amazon Macie (p. 53).

Activity ISPs

This view provides the complete list of CloudTrail data and management events monitored by Macie, grouped by the associated internet service providers (ISPs) for the specified user. For each ISP, the total count of the user sessions (5-minute integrations of CloudTrail data) that this ISP is present in, and the percentage that this total represents of the total number of user sessions is displayed.

To navigate from this view to the **Research** tab, choose the looking glass icon next to any of the displayed themes. Your selection is automatically translated into a query that appears in the query parser in the **Research** tab. For more information, see Researching Through Data Monitored by Amazon Macie (p. 53).

CloudTrail User Identity Types

This view provides the complete list of CloudTrail data and management events monitored by Macie, grouped by the user identity type for the specified users. For more information, see the definition for

user in Concepts and Terminology (p. 3) . For each user identity type, the total count of the user sessions (5-minute integrations of CloudTrail data) that this user identity type is present in, and the percentage that this total represents of the total number of user sessions is displayed.

To navigate from this view to the **Research** tab, choose the looking glass icon next to any of the displayed themes. Your selection is automatically translated into a query that appears in the query parser in the **Research** tab. For more information, see Researching Through Data Monitored by Amazon Macie (p. 53).

Researching Through Data Monitored by Amazon Macie

You can use the **Research** tab in the Macie console to construct and run queries in the query parser and conduct in-depth investigative research of your data and activity that Macie monitors. You can navigate to the **Research** tab at any time and construct queries in the empty parser. For more information, see Constructing Queries in Macie (p. 53). You can be redirected to the **Research** tab from various places throughout the Macie console: for example, any of the **Dashboard** views (see Using the Amazon Macie Dashboard (p. 34)) or the **Basic alerts** list (see Amazon Macie Alerts (p. 41)). When redirected to the **Research** tab from other places in the console, your data selection is translated into an automatically generated query that appears in the query parser.

Topics

Constructing Queries in Macie

Macie enables you to construct queries in the query parser in the **Research** tab. The query parser is a lexer that interprets a string into a Lucene Query using JavaCC. For more information about query syntax, see Apache Lucene - Query Parser Syntax.

The following are example queries for common searches:

- To search for any console login not that didn't originate from IP addresses owned by Amazon: `eventNameIsp.compound:/ConsoleLogin:~(Amazon.*)/`
- To search for PII artifacts inside a public S3 bucket: `filesystem_metadata.bucket:"my-public-bucket" AND (pii_impact:"moderate" OR pii_impact:"high")`

The following tables contains example queries for the Macie date, integer, and string field types.

Example Queries: Date Field Type

Example Query	Description	Data Repository
`objectsRead.key:* AND @timestamp:[2017-08-01 TO 2017-12-31]`	Search for S3 objects read in the fourth quarter of 2017.	CloudTrail data
`sourceIPAddress.ip_intel.type:"Tor" AND @timestamp:[now-1M TO now]`	Search for anonymous accesses to your Macie-monitored data	CloudTrail data

Example Query	Description	Data Repository
	from Tor exit notes over the last month.	
`macieUniqueId:"085924634393 AND role\:malicious_user" AND @timestamp: [2018-01-18 TO *]`	Search for AWS activities of an assumed role named "malicious_user" in the AWS account ID 085924634393, starting from January 18, 2018.	CloudTrail data

Example Queries: Integer Field Type

Example Query	Description	Data Repository
`dlp_risk>6 AND filesystem_metadata.server_encryption:""`	Search for S3 objects with a `dlp_risk` score greater than 6 and without a server-side encryption.	S3 objects
`filesystem_metadata.size: [10240 TO 1024000] AND pii_types:*`	Search for S3 objects between the sizes of 10 MB to 1 GB that contain potential PII data.	S3 objects

Example Queries: String Field Type

Example Query	Description	Data Repository
`dlp_risk>5 AND key: /.*contract.*|.*agreement.*|.*terms.*/ AND @timestamp:[now-1M/M TO now]`	Search for S3 object keys (names) that contain the keywords "contract," "agreement," or "terms," with a `dlp_risk` score higher than 5, and that were last modified less than a month ago. **Note** Some regex queries might result in long search times. We recommend conducting searches for limited time frames.	S3 objects
`mimetypes:"Adobe PDF \(application/pdf\)" AND key: /~(.*\.pdf|.*\.PDF)/`	Search for S3 objects containing PDF data but in files with file extensions other than PDF/pdf. **Note** This query also returns archived objects (zip,7z, etc.) containing PDF documents.	S3 objects

Example Query	Description	Data Repository
`acl.Grants.Grantee.DisplayName: admin`	Search for S3 buckets with ACL grantee display names set to "admin."	S3 bucket properties
`acl.Grants.Grantee.DisplayName: admi?`	Search for S3 buckets with ACL grantee display names set to "admi(?)" (wildcard), including "admin."	S3 bucket properties
`bucket: *test*`	Search for S3 buckets with keywords "test."	S3 bucket properties

Research Filters

In the Macie **Research** tab, you can apply the following filters to your searches.

Data Index

The first **Research** tab filter (dropdown) with the preselected default value of **CloudTrail data**, enables you to specifying the index (or the data repository) that you want Macie to search through. This filter includes the following options:

- **CloudTrail data** – A collection of 5-minute aggregates of raw CloudTrail data
- **S3 bucket properties** – A collection of metadata about the S3 buckets that Macie is monitoring
- **S3 objects** – A collection of metadata about the S3 objects that are stored in the buckets that Macie is monitoring

Number of Results to Display

The next **Research** tab filter with the preselected default value of **Top 10** enables you to control the number of results to display when you do your initial search and the number of additional results to display if more results are available. This filter includes the following options:

- Top 10
- Top 50
- Top 100
- Top 500

Time Range

The third **Research** tab filter with the preselected default value of **Past 30 days** enables you to define a time range that you want to display your search results for. This filter includes the following options:

- Past 7 days
- Past 30 days
- Past 90 days
- Past 365 days

- All
- Custom time range

Saving a Query as an Alert

You can use the following procedure to save a query that appears in the query parser as a basic alert. For more information about basic alerts, see Amazon Macie Alerts (p. 41).

1. In the Macie console's **Research** tab, either autogenerate or construct a query in the query parser.
2. Choose the **Save query as alert** icon.
3. Fill out the **Basic alert definition** form and choose **Save**. For more information, see Adding New and Editing Existing Custom Basic Alerts (p. 44).

Favorite Queries

You can mark queries that you frequently run as favorites and view a list of your favorite queries.

1. In the Macie console's **Research** tab, either autogenerate or construct a query in the query parser.
2. Choose the **Mark query as favorite** icon.
3. Fill out the **Favorite query definition** form by specifying the name and the description for the favorite query and choose **Save**.
4. To view the list of your favorite queries, in the Macie console's **Research** tab, choose the **Favorite queries** icon.

Researching AWS CloudTrail Data

Topics

Analyzing CloudTrail Search Results

The following section describes the elements of the search results that get displayed when you use the **Research** tab to investigate your Macie-monitored CloudTrail data.

Complete the following steps in the **Research** tab.

1. Select **CloudTrail data** in the first filter dropdown.
2. For this example, select **Top 10** in the second filter dropdown.
3. For this example, select **Past 90** days in the third filter dropdown.
4. Choose the button with the looking glass icon to start the search.

Your search produces the following elements:

- The **total number of results** that matched your CloudTrail data search for the selected time range.
- The **graphical representation** of CloudTrail data search results for the selected time range.

Note
If your dataset is very large and you specify a very wide time range, your data might not render properly, and this graph might not appear as one of the resulting elements of your search.

Important
You can use the graph to further narrow your search and generate and run a query that produces a subset of the results generated by your original selections in the preceding steps. Double-click any of the graph's results and your selection is translated into a new query that automatically appears in the query parser, and the **Research** tab is refreshed with the results of this new query.

- **Search results summary** – A list of the most significant fields from your search. The first line includes the top (or bottom) three values for each field. The second line includes the top (or bottom) 10 values for each field.

  Important
  You can use the fields in the search results summary to further narrow your search and generate and run a query that produces a subset of the results generated by your original selections in the preceding steps. Choose the first or the second line of results for any field, and in the expanded results breakdown, choose the looking glass icon next to any of the results. Your choices are then translated into a new query that automatically appears in the query parser, and the **Research** tab is refreshed with the results of this new query.
- A list of **user sessions** (5-minute aggregates of CloudTrail data) that match your search criteria. Choose any user session to expand it and view its details.

CloudTrail Data Fields and Sample Queries

The following tables include the fields that can appear in the results of your CloudTrail data searches.

- The first table includes the fields that Macie extracts from CloudTrail. These fields also include Amazon S3 data events. For example, `accountId` in Macie corresponds to `userIdentity.accountId` in CloudTrail, and `eventNameErrorCode.key` in Macie corresponds to `eventName` in CloudTrail.
- The second table includes the fields that Macie generates to provide further security intelligence and context based on the examined CloudTrail data. For example, `isp.key` describes the organization or the ISP that the API request against your AWS resources is coming from, and `sourceIPAddress.ip_intel.type` describes the IP address history: for example, whether it's a Tor exit node that is being used to initiate API requests against your AWS resources.

CloudTrail Data Fields That Macie Extracts

Note
For this data repository (CloudTrail), your search always returns a list of user sessions: 5-minute aggregates of raw CloudTrail data. A user session is determined by the Macie unique ID: a format that is unique to Macie for specifying users. Macie unique ID is a combination of the IAM `UserIdentity` element and the `recipientAccountId`.

Macie Field Name	CloudTrail Field Name	Macie Field Type	Description	Example Search Query
`accountId`	`userIdentity.accountId`	String	The AWS account ID.	Search for user sessions with accesses related to a particular account: • **accountId:"110912345678"**

Macie Field Name	CloudTrail Field Name	Macie Field Type	Description	Example Search Query
`awsRegion.key`	`awsRegion`	String	The AWS Region that the request is made to.	Search for user sessions with any AWS API calls by Region: • **awsRegion.key:"us-west-2"** • **awsRegion.key:"us-east-1"**
`eventNameErrorCode.key`	`eventName` and `errorCode`	String	The event name that resulted in the returned (if any) error code.	• Search for user sessions with any AWS `ConsoleLogin` call: • **eventNameErrorCode.key:Consol** • Search for user sessions with any AWS `Delete` call: • **eventNameErrorCode.key:Delete**
`eventNameErrorCode.secondary`	`errorCode`	String	The error code returned after an unsuccessful API request.	Search for user sessions with any `AccessDenied` error across all CloudTrail API events: • **eventNameErrorCode.secondary:"A**
`eventSource.key`	`eventSource`	String	The service that the request was made to.	Search for user sessions with any API calls of a particular AWS service: • **eventSource.key:"s3.amazonaws.co** • **eventSource.key:"lambda.amazona**
`eventType.key`	`eventType`	String	The type of the event that generated the event record (for example, `AwsApiCall`, `AwsServiceEvent`, or `AwsConsoleSignIn`).	Search for user sessions with any AWS API calls of a particular `eventType`: • **eventType.key:"AwsApiCall"**

Macie Field Name	CloudTrail Field Name	Macie Field Type	Description	Example Search Query
`objectsDeleted.key`	`resources[0].ARN`	String	A list of S3 objects ARNs, S3 bucket ARNs, or prefix ARNs that were part of a `DeleteObject` or `DeleteObjects` API call. **Note** When you delete an S3 bucket, both `DeleteBucket` and `DeleteObjects` APIs are called. The aggregate record with the `DeleteObjects` call lists the deleted bucket or prefix, not all the individual objects that were deleted. **Note** Objects that are part of a failed `DeleteObject` or `DeleteObjects` API call are also added to the aggregate record of `objectsDeleted.key`. **Note** A user session returning the results of a search against `objectsDeleted.key` has a maximum limit of 250 records.	Search for all objects deleted from a particular bucket or prefix: • **objectsDeleted.key:/arn:aws:s3:::my_sensitive_bucket.*/** Search for all `Delete` requests of a particular object that were made anonymously or by any user or role. • **objectsDeleted.key: "arn:aws:s3:::my-bucket-name/sshKeys"** Search for user sessions that contain both a `DeleteObject:AccessDenied` and any attempt to delete a particular sensitive object, bucket, or prefix. • **objectsDeleted.key:/arn:aws:s3:::my_sensitive_bucket.*/ AND eventNameErrorCode.compound:"** Search for user sessions that contain both an attempt (or attempts) to delete S3 objects from outside AWS and any attempt to delete a particular sensitive object, bucket, or prefix: • **objectsDeleted.key:/arn:aws:s3:::my_sensitive_bucket.*/ AND eventNameIsp.compound:/DeleteObject:~(Amazon.*)/** Search for anonymous delete requests of a known sensitive object:

Macie Field Name	CloudTrail Field Name	Macie Field Type	Description	Example Search Query
				• **objectsDeleted.key:/ arn:aws:s3:::my_sensitive_bucket.* AND accountId:"ANONYMOUS_PRINCIP**

Macie Field Name	CloudTrail Field Name	Macie Field Type	Description	Example Search Query
`objectsRead.key`	`resources[0].ARN`	String	A list of S3 object ARNs that were part of a `GetObject` API call. **Note** Objects that are part of a failed `GetObject` API call are also added to the aggregate record of `objectsRead.key`. **Note** A user session that returns the results of a search against `objectsRead.key` has a maximum limit of 250 records.	Search for user sessions with all objects read from a particular bucket or prefix: • **objectsRead.key:/ arn:aws:s3:::my_sensitive_bucket.*/** Search for all access attempts of a particular object made either anonymously or by any user or role. • **objectsRead.key:"arn:aws:s3:::my-bucket-name/ sshKeys"** Search for user sessions that contain both a `GetObject:AccessDenied` and any attempt to read a particular sensitive object, bucket, or prefix. • **objectsRead.key:/ arn:aws:s3:::my_sensitive_bucket.*/ AND eventNameErrorCode.compound:"** Search for user sessions that contain both an attempt (or attempts) to read S3 objects from outside AWS and any attempt to read a particular sensitive object, bucket, or prefix: • **objectsRead.key:/ arn:aws:s3:::my_sensitive_bucket.*/ AND eventNameIsp.compound:/ GetObject:~(Amazon.*)/** Search for anonymous read accesses to a known sensitive object or bucket:

Macie Field Name	CloudTrail Field Name	Macie Field Type	Description	Example Search Query
				• **objectsRead.key:/ arn:aws:s3:::my_sensitive_bucket.*/ AND accountId:"ANONYMOUS_PRINCIPA**

Macie Field Name	CloudTrail Field Name	Macie Field Type	Description	Example Search Query
`objectsWritten.key`	`resources[0].ARN`	String	A list of S3 object ARNs that were part of a `PutObject`, `CopyObject`, or `CompleteMultipartUpload` API call. **Note** Objects that are part of a failed `PutObject` API call are also added to the aggregate record of `objectsWritten.key`. **Note** A user session that returns the results of a search against `objectsWritten.key` has a maximum limit of 250 records.	Search for user sessions with all objects written to a particular bucket: • **objectsWritten.key:/arn:aws:s3:::my_bucket_name.*/** Search for user sessions with all write requests of a particular object made either anonymously or by any user or role: • **objectsWritten.key:"arn:aws:s3:::my-bucket-name/sshKeys"** Search for user sessions that contain both a `PutObject:AccessDenied` and any attempt to read a particular sensitive object, bucket, or prefix. • **objectsWritten.key:/arn:aws:s3:::my_sensitive_bucket.*/ AND eventNameErrorCode.compound:"** Search for user sessions that contain both an attempt (or attempts) to write S3 objects from outside AWS and any attempt to write a particular sensitive object, bucket, or prefix: • **objectsWritten.key:/arn:aws:s3:::my_sensitive_bucket.*/ AND eventNameIsp.compound:/PutObject:~(Amazon.*)/** Search for anonymous write requests to a sensitive object or bucket:

Macie Field Name	CloudTrail Field Name	Macie Field Type	Description	Example Search Query
				• **objectsWritten.key:/ arn:aws:s3:::my_sensitive_bucket.*/ AND accountId:"ANONYMOUS_PRINCIPA**
`principalId`	`userIdentity.principalId`	String	The IAM principal ID. **Note** When an assumed role makes a request, the session name is removed from the principal ID.	Search for user sessions with access requests from a particular principal ID: • **principalId:"AIDAIMABCKFJSKEOAH**
`recipientAccountId`	`recipientAccountId`	String	The account ID that received the CloudTrail event.	Search for all activity in a particular account: • **recipientAccountId:"110912345678** Search for anonymous access requests to a particular account: • **recipientAccountId:"110912345678 AND accountId: "ANONYMOUS_PRINCIPAL"**
`resourceOwnerAccountIds.key`	`Resources[].accountId`	String	List of AWS resource owners. An example is a list of account IDs that own an S3 object or bucket.	Search for activity against resources owned by a particular account: • **resourceOwnerAccountIds.key: "110951234567"**
`resources.key`	`Resources[0].arn`	String	List of resources (S3 buckets only) associated with the CloudTrail events in the user session.	Search for access requests to a particular S3 bucket: • **resources.key: "arn:aws:s3:::my-bucket-name"** Search for anonymous access requests to a known sensitive bucket: • **resources.key: "arn:aws:s3:::my-super-sensitive-bucket" AND accountId:"ANONYMOUS_PRINCIPA**

Macie Field Name	CloudTrail Field Name	Macie Field Type	Description	Example Search Query
`sessionName.key`	`userIdentity.principalId`	String	The identifier for the assumed role session. When an assumed role makes a request, the session name is removed from the principal ID and is assigned as a value to `sessionName.key`. When an identity other than an assumed role makes a request, `sessionName.key` is set to `None`.	Search for assumed role access requests from session name `examplesession-cli`: • **sessionName.key:"examplesession-cli"** Search for EC2 instance IDs in session names: • **(sessionName.key:/i-[0-9a-f]{8}/ OR sessionName.key:/i-[0-9a-f]{17}/)** Search for assumed role access requests to a role from a `sessionName` other than `examplesession-cli` using regex negation: • **macieUniqueId: "123456789123:assumed-role:co-admin" AND sessionName.key:/~(examplesession-cli)/**
`sourceARN`	`userIdentity.arn`	String	The ARN used to make the request. **Note** When an assumed role makes a request, the session name is removed from `sourceARN`.	Search for user sessions with access requests from a particular ARN: • **sourceARN:"arn:aws:iam::1234567 cluster-api"**

Macie Field Name	CloudTrail Field Name	Macie Field Type	Description	Example Search Query
`sourceIPAddress.key`	`sourceIPAddress`	String	The IP address that the request was made from. **Note** A user session that returns the results of a search against `sourceIPAddress.key` has a maximum limit of 60,000 records.	Search for user sessions with access requests from a particular source IP address: • **sourceIPAddress.key:"194.68.22.22** Search through user sessions with source IP addresses using wildcards: • **sourceIPAddress.key:194.68.*.*** Search for user sessions with more than 10 `RunInstances` events and without any events requested by the `autoscaling` group: • **eventNameErrorCode.RunInstance AND NOT (sourceIPAddress.key:"autoscaling.**
`userAgent.key`	`userAgent`	String	A list of client user agent strings used to make the AWS API call.	Search for user sessions with API calls executed by Amazon S3: • **userAgent.key:"s3.amazonaws.com**
`userIdentityType.key`	`userIdentity.type`	String	A list of identity types in AWS.	Search for user sessions with access requests by the root identity in an account: • **userIdentityType.key:"Root"**

Fields That Macie Generates

Note
For this data repository (CloudTrail), your search always returns a list of user sessions: 5-minute aggregates of raw CloudTrail data. A user session is determined by the Macie unique ID: a format that is unique to Macie for specifying users. The Macie unique ID is a combination of the IAM `UserIdentity` element and the `recipientAccountId`.

Macie Field Name	Macie Field Type	Description	Example Search Query
`@timestamp`	Date	The start time of a user session.	Search for user sessions with access requests after a specific time: • **@timestamp:>"2017-02-06T23:01:08Z"**

Macie Field Name	Macie Field Type	Description	Example Search Query
			• **@timestamp:>"2017-02-06"** Search for user sessions with access requests between two time intervals: • **@timestamp:[2017-02-01 TO 2017-02-27]**
`countLongLifeAccessToken`	Integer	A count of `GetSessionToken` API calls with a lifespan longer than the default 43,200 seconds.	Search for user sessions with a user or role creating a temporary access token with a longer than the default lifespan: • **countLongLifeAccessToken:>0**
`dcObjectsDeleted`	Integer	A count of unique S3 objects deleted in a user session. **Note** A user session that returns the results of a search against `dcObjectsDeleted` has a maximum limit of 250 entries.	Search for user sessions with more than 25 distinct objects deleted by an AWS user or a role: • **dcObjectsDeleted:>25** • **dcObjectsDeleted:[25 TO 100]**
`dcObjectsRead`	Integer	A count of unique S3 objects read in a user session. **Note** A user session that returns the results of a search against `dcObjectsRead` has a maximum limit of 250 entries.	Search for user sessions with more than 25 distinct objects read by an AWS user or a role: • **dcObjectsRead:>25** • **dcObjectsRead:[25 TO 100]** Search for more than 25 distinct objects read by an anonymous principal during a user session: • **dcObjectsRead:>25 AND accountId:"ANONYMOUS_PRINCIPAL"**
`dcObjectsWritten`	Integer	A count of unique S3 objects written in a user session. **Note** A user session that returns the results of a search against `dcObjectsWritten` has a maximum limit of 250 entries.	Search for user sessions with more than 25 distinct objects written by an AWS user or a role: • **dcObjectsWritten:>25** • **dcObjectsWritten:[25 TO 100]**

Macie Field Name	Macie Field Type	Description	Example Search Query
`distinctEventName`	Integer	A count of unique event names that take place in a user session.	Search for user sessions with more than 25 unique API calls being executed by a user or a role: • **distinctEventName:>25** • **distinctEventName:[25 TO 100]**
`distinctSourceIPAddress`	Integer	A count of unique source IP addresses involved in activity that takes place in a user session. The maximum value is 60,000.	Search for user sessions with more than 25 distinct source IP addresses observed for a user or a role: • **distinctSourceIPAddress:>25** • **distinctSourceIPAddress: [25 TO 100]**
`distinctUserAgent`	Integer	A count of unique client user agents involved in activity that takes place in a user session. The maximum value is 60,000.	Search for user sessions with more than 25 user agents observed for a user or a role: • **distinctUserAgent:>25** • **distinctUserAgent:[25 TO 100]**
`eventNameErrorCode.compound`	String	A compound aggregation that summarizes each CloudTrail event name along with any error codes that are associated with the API Call. The format is `EventName:ErrorCode` for the term value, which enables Macie to associate an API event name with the error code, if any, that is returned. If there is no error code for the event, the value is set only to the API name with no colon, for example: `PutObject`.	Search for user sessions with `AccessDenied` error while attempting a `GetObject` call: • **eventNameErrorCode.compound:"GetO** Search for user sessions with any errors associated with `PutObject` calls: • **eventNameErrorCode.compound:/PutObject:.*/**
`eventNameIsp.compound`	String	A compound aggregation that summarizes each CloudTrail event name along with the Internet Service Provider (ISP) that the request originated from. The format is `EventName:ISP` for the term value, which enables Macie to associate an API operation name with the ISP that it originated from.	Search for user sessions with `ConsoleLogin` calls from non-AWS IPs using a regular expression: • **eventNameIsp.compound:/ConsoleLogin:~(Amazon.*)/**

Macie Field Name	Macie Field Type	Description	Example Search Query
`eventNameIsp.secondary`	String	The ISP that the AWS API call was made from.	Search for user sessions with AWS API calls coming from outside Amazon IP addresses: • **eventNameIsp.secondary:/~(Amazon.*)/**
`macieUniqueId`	String	A format that is unique to Macie for specifying users. The Macie unique ID is a combination of the IAM `UserIdentity` element and the `recipientAccountId`. For more information, see MacieUniqueID (p. 48).	Search for user sessions with accesses from a particular role, user, or root account: • **macieUniqueId:"123456789123:assume** **role:co-admin"** • **macieUniqueId:"123456789123:root:roo** • **macieUniqueId:"123456789123:user:exa**
`sourceIPAddress.ip_intel.type`	String	The IP intelligence category associated with a source IP address.	Search for user sessions with all accesses from a Tor network: • **sourceIPAddress.ip_intel.type:"TOR"** Search for user sessions with all accesses from threat intelligence input feeds: • **sourceIPAddress.ip_intel.type:***
`windowStartTimeInMillis`	Integer	The epoch timestamp for the start of a user session.	Search for user sessions whose first event time is greater than a given epoch time: • **windowStartTimeInMillis:>1424476529**
`windowEndTimeInMillis`	Integer	The epoch timestamp for the end of a user session.	Search for user sessions whose last event time is less than a given epoch time: • **windowEndTimeInMillis:<1424476987**
`ipLocation.key`	String	The IP geolocation (city and country) accessed by an identity that Macie monitors.	Search for user sessions with any AWS API call events originating in Los Angeles: • **ipLocation.key:"LosAngeles:UnitedState** Search for user session any AWS API call events originating from outside the United States: • **ipLocation.key:/~(.*UnitedStates)/**

Macie Field Name	Macie Field Type	Description	Example Search Query
`isp.key`	String	The ISP that the AWS API call originated from.	Search for user sessions with AWS API calls coming from outside Amazon IP addresses: • **isp.key:/~(Amazon.*)/**

Researching S3 Bucket Properties Data

Topics

Analyzing S3 Buckets Properties Search Results

The following section describes the elements of the search results that appear when you use the **Research** tab to investigate your S3 bucket properties data that Macie monitors.

Complete the following steps in the **Research** tab.

1. Select **S3 bucket properties** in the first filter dropdown.
2. For this example, select **Top 10** in the second filter dropdown.
3. For this example, select **Past 90** days in the third filter dropdown.
4. Choose the button with the looking glass icon to start the search.

Your search results contain the following elements:

- The **total number of results** that matched your S3 bucket properties data search for the selected time range.
- The **graphical representation** of the S3 bucket properties data search results for the selected time range.

  Note
  If your dataset is very large and you specify a very wide time range, your data might not render properly, and this graph might not appear as one of the resulting elements of your search.

  Important
  You can use the graph to further narrow your search and generate and run a query that produces a subset of the results generated by your original selections in the preceding steps. Double-click any of the graph's results, and your selection is translated into a new query that automatically appears in the query parser, and the **Research** tab is refreshed with the results of this new query.
- **Search results summary** – A list of the most significant fields from your search. The first line includes the top (or bottom) three values for each field. The second line includes the top (or bottom) 10 values for each field.

  Important
  You can use the fields in the search results summary to further narrow your search and generate and run a query that produces a subset of the results generated by your original selections in the preceding steps. Choose the first or the second line of results for any field,

and in the expanded results breakdown, choose the looking glass icon next to any of the results. Your choices are translated into a new query that automatically appears in the query parser, and the **Research** tab is refreshed with the results of this new query.

- A list of S3 buckets that match your search criteria. Choose any bucket to expand it and view its details.

S3 Bucket Properties Data Fields and Example Queries

The following tables include the fields that can appear in the results of your S3 buckets metadata searches:

- The first table includes the fields that Macie extracts from the Amazon S3 bucket API metadata. For example, `acl.Grants.Grantee.DisplayName` in Macie corresponds to `Grants.Grantee.DisplayName` in the Amazon S3 `getbucket-acl` API response.
- The second table includes the fields that Macie generates to provide further security intelligence and context based on the examined S3 buckets metadata. For example, `s3_world_readability` describes a true/false/unknown state condition of whether an S3 bucket is readable by everyone as part of evaluating its Amazon S3 ACL and bucket (IAM) policy.

S3 Bucket Properties Data Fields That Macie Extracts

Macie Field Name	Amazon S3 API Field Name	Amazon S3 API Operation	Macie Field Type	Description	Example Search Query
`acl.Grants.Grantee.DisplayName`	`Grantee.DisplayName`	`bucket-acl`	String	The display name of the S3 bucket ACL grantee.	Search for S3 buckets accessible by John Doe: • **acl.Grants.Grantee.DisplayNam**
`acl.Grants.Grantee.ID`	`Grantee.ID`	`bucket-acl`	String	The ID of the identity that was granted access to the S3 bucket by the bucket owner.	Search for an S3 bucket's grantee with a particular canonical ID: • **acl.Grants.Grantee.ID:"75bee88**
`acl.Grants.Grantee.Type`	`Grantee.Type`	`bucket-acl`	String	The user type of the S3 bucket ACL grantee.	Search for all S3 buckets that are granted to `Users`: • **acl.Grants.Grantee.Type:Canoni** Search for all S3 buckets that are granted to `Groups`: • **acl.Grants.Grantee.Type:Group**
`acl.Grants.Grantee.URI`	`Grantee.URI`	`bucket-acl`	String	The URI identifier of the S3 bucket ACL grantee.	Search for all S3 buckets except those that

Macie Field Name	Amazon S3 API Field Name	Amazon S3 API Operation	Macie Field Type	Description	Example Search Query
					belong to the `LogDelivery` group: • **acl.Grants.Grantee.URI:"http://acs.amazonaws.com/groups/s3/LogDelivery"** Search for all S3 buckets that have global share permissions: • **acl.Grants.Grantee.URI:"http://acs.amazonaws.com/groups/global/AllUsers"** Search for all S3 buckets that allow access to (any) AWS authenticated users: • **acl.Grants.Grantee.URI:"http://acs.amazonaws.com/groups/global/AuthenticatedUsers"**
`acl.Grants.Permission`	`Grants.Permission`	`get-bucket-acl`	String	The permission level assigned to the ACL grantee.	Search for S3 buckets that grant full (read/write) access to anyone: • **acl.Grants.Grantee.URI:"http://acs.amazonaws.com/groups/global/AllUsers" AND acl.Grants.Permission:"FULL_CO**
`acl.Owner.DisplayName`	`Owner.DisplayName`	`get-bucket-acl`	String	The display name of the S3 bucket owner.	Search for S3 buckets owned by John Doe: • **acl.Owner.DisplayName: "JohnDoe"**
`acl.Owner.ID`	`Owner.ID`	`get-bucket-acl`	String	The ID of the S3 bucket owner.	Search for a particular S3 bucket owner ID: • **acl.Owner.ID:"73bee78dfe7b89**

Macie Field Name	Amazon S3 API Field Name	Amazon S3 API Operation	Macie Field Type	Description	Example Search Query
location.LocationConstraint	LocationConstraint	get-bucket-location	String	The AWS Region where the S3 bucket resides. **Note** By default, buckets in the us-east-1 Region have no region returned from the S3 API call. To facilitate searching, Macie automatically populates them with the string "us-east-1".	Search for buckets hosted in the us-west-2 Region: • **location.LocationConstraint:"us-west-2"** Search for buckets hosted in the us-east-1 Region: • **location.LocationConstraint:"us-east-1"**
logging.LoggingEnabled.TargetBucket	LoggingEnabled.TargetBucket	get-bucket-logging	String	The bucket whose logging status is being returned.	Search for all buckets with S3 object level logging enabled: • **logging.LoggingEnabled.Target**
logging.LoggingEnabled.TargetPrefix	LoggingEnabled.TargetPrefix	get-bucket-logging	String	The configured prefix or folder containing Object Level Logging data for a particular S3 bucket.	Search for buckets configured with a prefix substring of "Production": • **logging.LoggingEnabled.Target "Production"**
policy.Policy.Id	Policy.Id	get-bucket-policy	String	The ID for an S3 bucket policy.	Search for bucket policies with a particular ID: • **policy.Policy.Id:"aaaa-bbbb-cccc-dddd"**

Macie Field Name	Amazon S3 API Field Name	Amazon S3 API Operation	Macie Field Type	Description	Example Search Query
policy.Policy.Statement.Action	Policy.Statement.Action	GetBucketPolicy	String	The list of actions (API requests) associated with an S3 bucket policy.	Search for bucket policies with "put" substring actions (PutObject, PubBucketPolicy, etc.): • **policy.Policy.Statement.Action: s3:Put.*/**
policy.Policy.Statement.Effect	Policy.Statement.Effect	GetBucketPolicy	String	The list of policy effects associated with an S3 bucket policy.	Search for bucket policies with explicit "allow" grants: • **policy.Policy.Statement.Effect: "Allow"**
policy.Policy.Statement.NotPrincipal.AWS	Policy.Statement.NotPrincipal.AWS	GetBucketPolicy	String	The principal exception to which the policy rule is applied.	Search for bucket policies with a particular account specified in the NotPrincipal section: • **policy.Policy.Statement.NotPri account-ID:role/ role-name"**
policy.Policy.Statement.NotPrincipal.CanonicalUser	Policy.Statement.NotPrincipal.CanonicalUser	GetBucketPolicy	String	The CanonicalUser stated in the NotPrincipal expression of the policy.	Search for bucket policies with a particular CanonicalUser specified in the NotPrincipal section: • **policy.Policy.Statement.NotPri**
policy.Policy.Statement.NotPrincipal.Federated	Policy.Statement.NotPrincipal.Federated	GetBucketPolicy	String	The Federated identity stated in the NotPrincipal expression of the policy.	Search for bucket policies with a particular Federated user specified in the NotPrincipal section: • **policy.Policy.Statement.NotPri account-ID:saml-provider/ provider-name"**

Macie Field Name	Amazon S3 API Field Name	Amazon S3 API Operation	Macie Field Type	Description	Example Search Query
policy.Policy.Statement.NotPrincipal.Service	Policy.Statement.NotPrincipal.Service	GetBucketPolicy, bucket-policy	String	The Service stated in the NotPrincipal expression of the policy.	Search for bucket policies with a particular Service specified in the NotPrincipal section: • **policy.Policy.Statement.NotPri**
policy.Policy.Statement.Principal.AWS	Policy.Statement.Principal.AWS	GetBucketPolicy, bucket-policy	String	The principal specified in the AWS expression.	Search for bucket policies with explicit allow grants to any AWS resource: • **policy.Policy.Statement.Effect: Allow AND policy.Policy.Statement.Princip \"*\"**
policy.Policy.Statement.Principal.CanonicalUser	Policy.Statement.Principal.CanonicalUser	GetBucketPolicy, bucket-policy	String	The CanonicalUser stated in the principal expression of the policy.	Search for bucket policies with a particular CanonicalUser specified in the Principal section: • **policy.Policy.Statement.Princip**
policy.Policy.Statement.Principal.Federated	Policy.Statement.Principal.Federated	GetBucketPolicy, bucket-policy	String	The Federated identity stated in the principal expression of the policy.	Search for bucket policies with a particular Federated user specified in the NotPrincipal section: • **policy.Policy.Statement.NotPri account-ID:saml-provider/ provider-name"**
policy.Policy.Statement.Principal.Service	Policy.Statement.Principal.Service	GetBucketPolicy, bucket-policy	String	The Service stated in the principal expression of the policy.	Search for bucket policies with a particular Service user specified in the NotPrincipal section: • **policy.Policy.Statement.NotPri**

Macie Field Name	Amazon S3 API Field Name	Amazon S3 API Operation	Macie Field Type	Description	Example Search Query
policy.Policy.Statement.Resource	Policy.Statement.Resource	get-bucket-policy	String	The S3 resource that the S3 bucket policy is applied to.	Search for S3 bucket policies containing wildcards: • **policy.Policy.Statement.Resour */**
policy.Policy.Statement.Sid	Policy.Statement.Sid	get-bucket-policy	String	The `Sid` of the S3 bucket policy.	Search for bucket policies with a particular `Sid`: • **policy.Policy.Statement.Sid:"1"**
policy.Policy.Version	Policy.Version	get-bucket-policy	String	The version number for the S3 bucket policy.	Search for bucket policies with a particular version: • **policy.Policy.Statement.Version**
tagging.TagSet.Key	TagSet.Key	get-bucket-tagging	String	The key of the S3 bucket tag.	Search for bucket policies with a particular tag key: • **tagging.TagSet.Key:"User"**
tagging.TagSet.Value	TagSet.Value	get-bucket-tagging	String	The value of the S3 bucket tag.	Search for bucket policies with a particular tag value: • **tagging.TagSet.Value:"johndoe"**
versioning.MFADelete	MFADelete	get-bucket-versioning	String	The `MFADelete` (enabled/disabled) state of the bucket version configuration.	Search for buckets where `MFADelete` is enabled in the bucket versioning configuration: • **versioning.MFADelete:"enabled**
website.ErrorDocument.Key	ErrorDocument.Key	get-bucket-website	String	The error document configured as part of S3 static website hosting.	Search for S3 buckets configured for static website hosting and with an error page redirection to 404.html: • **website.ErrorDocument.Key: "404.html"**

Macie Field Name	Amazon S3 API Field Name	Amazon S3 API Operation	Macie Field Type	Description	Example Search Query
website.IndexDocument.Suffix	IndexDocument.Suffix	bucket-website	String	The suffix of a webpage that Amazon S3 returns when a request is made to the root of a website or any subfolder.	Search for the index document configured as part of S3 static website hosting and with an index page redirection to index.html: • **website.IndexDocument.Key: "index.html"**
• lifecycle_configuration.Rules.Expiration.Date • lifecycle_configuration.Rules.Expiration.Days • lifecycle_configuration.Rules.AbortIncompleteMultipartUpload.DaysAfterInitiation • lifecycle_configuration.Rules.Filter.Prefix • lifecycle_configuration.Rules.Filter.Tag.Key • lifecycle_configuration.Rules.Filter.Tag.Value • lifecycle_configuration.Rules.ID • lifecycle_configuration.Rules.NoncurrentVersionExpiration.NoncurrentDays • lifecycle_configuration.Rules.NoncurrentVersionTransitions.NoncurrentDays • lifecycle_configuration.Rules.NoncurrentVersionTransitions.StorageClass • lifecycle_configuration.Rules.Prefix • lifecycle_configuration.Rules.Status • lifecycle_configuration.Rules.Transitions.Date • lifecycle_configuration.Rules.Transitions.Days • lifecycle_configuration.Rules.Transitions.StorageClass	• Rules.Expiration.Date • Rules.Expiration.Days • Rules.AbortIncompleteMultipartUpload.DaysAfterInitiation • Rules.Filter.Prefix • Rules.Filter.And.Tag.Key • Rules.Filter.And.Tag.Value • Rules.ID • Rules.NoncurrentVersionExpiration.NoncurrentDays • Rules.NoncurrentVersionTransitions.NoncurrentDays • Rules.NoncurrentVersionTransitions.StorageClass • Rules.Prefix • Rules.Status • Rules.Transitions.Date • Rules.Transitions.Days • Rules.Transitions.StorageClass	bucket-lifecycle-configuration	• Date • Integer • String • String • String • String • Integer • Integer • String • String • String • Date • Integer • String	For more information, see GET Bucket lifecycle.	Search for the S3 buckets with a lifecycle configuration rule whose expiration is less than 3 days: • **lifecycle_configuration.Rules.Ex**

S3 Bucket Properties Data Fields That Macie Generates

Macie Field Name	Macie Field Type	Description	Example search query
@timestamp	Date	The timestamp when Macie last analyzed the bucket.	Search for S3 buckets that Macie analyzed in the last 24 hours: • **@timestamp:[now-1d TO now]**
accountId	String	The account ID of the S3 bucket owner.	Search for any S3 buckets that don't belong to a given account: • **NOT accountId: 110912345678**
bucket	String	The name of an S3 bucket.	Search for a particular S3 bucket by name:

Macie Field Name	Macie Field Type	Description	Example search query
			• **bucket: "MyBucket"**
`s3_world_readability`	String	A value indicating whether the S3 bucket is globally readable: `true`, `false`, or `unknown`. The `unknown` value indicates that Macie can't determine whether the S3 bucket is globally readable.	Search for S3 buckets that are globally readable by either the Amazon S3 ACL or bucket (IAM) policy: • **s3_world_readability: "true"**
`s3_world_writability`	String	A value indicating if the S3 bucket is globally writable: `true`, `false`, or `unknown`. The `unknown` value indicates that Macie can't determine whether the S3 bucket is globally writable.	Search for S3 buckets that is globally writable by either the Amazon S3 ACL or the bucket (IAM) policy: • **s3_world_writability: "true"**

Researching S3 Objects Data

Topics

Analyzing S3 Objects Search Results

The following section describes the elements of the search results that get displayed when you use the **Research** tab to investigate your S3 objects that Macie monitors.

Complete the following steps in the **Research** tab.

1. Select **S3 objects** in the first filter pull-down list.
2. For this sample procedure, select **Top 10** in the second filter dropdown.
3. For this sample procedure, select **Past 90** days in the third filter dropdown.
4. Choose the button with the looking glass icon to start the search.

Your search results include the following elements:

- The **total number of results** that matched your S3 objects search for the selected time range.
- The **graphical representation** of the S3 objects search results for the selected time range.

  Note
  If your dataset is very large and you specify a very wide time range, your data might not render properly, and this graph might not appear as one of the resulting elements of your search.

  Important
  You can use the graph to further narrow your search and generate and run a query that produces a subset of the results generated by your original selections in the preceding steps. Double-click any of the graph's results, and your selection is translated into a new query that

automatically appears in the query parser, and the **Research** tab is refreshed with the results of this new query.

- **Search results summary** – A list of the most significant fields from your search. The first line includes the top (or bottom) three values for each field. The second line includes the top (or bottom) 10 values for each field.

  Important
  You can use the fields in the search results summary to further narrow your search and generate and run a query that produces a subset of the results generated by your original selections in the preceding steps. Choose the first or the second line of results for any field, and in the expanded results breakdown, choose the looking glass icon next to any of the results. Your choices are translated into a new query that automatically appears in the query parser, and the **Research** tab is refreshed with the results of this new query.
- A list of S3 objects that match your search criteria. Choose any S3 object to expand it and view its details.

S3 Objects Data Fields and Sample Queries

The following tables include the fields that can appear in the results of your S3 object searches:

- The first table includes the fields that Macie extracts from the Amazon S3 object API metadata. These are Macie fields that are also found in S3 API metadata. For example, `filesystem_metadata.ETag` describes the entity tag of an S3 object based on the checksum or hash of its content.
- The second table includes the fields that Macie generates to provide further security intelligence and context based on the examined S3 objects content and metadata. For example, `dlp_risk` represents a weighted score describing the risk profile of an S3 object metadata and its content, and `pii_types` describes any personal identifiable information contained in an S3 object.

S3 Object Data Fields That Macie Extracts

Macie Field Name	Amazon S3 API Field Name	Amazon S3 API Operation	Macie Field Type	Description	Example Search Query
`key`	`key`	`get-bucket` (`listObjects`)	String	The S3 object key path.	Search for document names with the keyword "myobject": • **key: /.*myobject.*/**
`accountId`	None	None	String	The AWS account ID that owns the S3 object.	Search for S3 objects owned by a particular account ID: • **accountId:"110912345678"**
`filesystem_metadata.bucket`	None	None	String	The S3 bucket name that holds the S3 object.	Search for S3 objects in a particular S3 bucket: • **filesystem_metadata.bucket:"M**

Macie Field Name	Amazon S3 API Field Name	Amazon S3 API Operation	Macie Field Type	Description	Example Search Query
`filesystem_metadata.first_prefix`	None **Note** Macie uses the S3 key field and parses out everything before the first '/', not including the bucket name.	get-bucket (`listObjects`)	String	The name of the first folder that contains the S3 object.	Search for S3 objects contained in first folder names where folder name is `AWSLogs`: • **filesystem_metadata.first_prefi**
`filesystem_metadata.ETag`	`ETag`	get-bucket (`listBuckets`)	String	The entity tag as defined in RFC 2616.	Search for a particular `eTag`: • **filesystem_metadata.ETag:""8b**
`filesystem_metadata.bucket_owner.id`	`ID`	get-bucket-acl	String	The unique ID of the S3 bucket owner.	Search for S3 objects belonging to a particular owner ID: • **filesystem_metadata.bucket_ov "447fba12b05da301df359096f**
`filesystem_metadata.bucket_owner.name`	`DisplayName`	get-bucket-acl	String	The name of the S3 bucket owner.	Search for S3 objects owned by John Doe: • **filesystem_metadata.bucket_ov "JohnDoe"**
`filesystem_metadata.last_modified`	`LastModified`	get-bucket (`list-buckets`)	Date	The timestamp when the S3 object was last modified.	Search for S3 objects that were modified in the last 24 hours: • **filesystem_metadata.last_modi [now-1d TO now]**

Macie Field Name	Amazon S3 API Field Name	Amazon S3 API Operation	Macie Field Type	Description	Example Search Query
filesystem_metadata.server_encryption	GetObject.ServerSideEncryption	get-object	String	The server side encryption used to encrypt an S3 object.	Search for objects that aren't encrypted with the AES256 standard: • **NOT filesystem.metadata.server_enc "AES256"**
filesystem_metadata.size	GetObject.ContentLength	get-bucket (list-buckets)	Integer	The size of the S3 object's content in bytes.	Search for S3 objects that are larger than 1 MB: • **filesystem.metadata.size: > 1024000**
filesystem_metadata.sse_kms_key_id	GetObject.SSEKMSKeyId	get-object	String	The unique identifier (ARN) of the master key used for server side encryption of the S3 objects.	Search for all S3 objects encrypted with a given key ID: • **filesystem_metadata.sse_kms_ "arn:aws:kms:us-west-2: 110912345678:key/06f8b4fa-3 b60a56a9a1f2"**
object_acl.Grants.Grantee.DisplayName	GetObjectAcl.Grants.Grantee.DisplayName	get-object-acl	String	The ACL grantee name.	Search for S3 object ACL permissions granted to John Doe: • **object_acl.Grants.Grantee.Displ "JohnDoe"**
object_acl.Grants.Grantee.ID	GetObjectAcl.Grants.Grantee.ID	get-object-acl	String	The ACL grantee unique ID.	Search for S3 object ACL permissions with a particular grantee ID: • **object_acl.Grants.Grantee.ID:"7**
object_acl.Grants.Grantee.Type	GetObjectAcl.Grants.Grantee.Type	get-object-acl	String	The ACL grantee type, such as `CanonicalUser` or `Group`.	Search for all S3 object ACLs that are granted to users or groups: • **object_acl.Grants.Grantee.Type** • **object_acl.Grants.Grantee.Type**

Macie Field Name	Amazon S3 API Field Name	Amazon S3 API Operation	Macie Field Type	Description	Example Search Query
`object_acl.Grants.Grantee.URI`	Grants.Grantee.URI	`get-object-acl`	String	The ACL grantee URI.	Search for S3 object ACLs with the `AllUsers` grant: • **object_acl.Grants.Grantee.URI: "http://acs.amazonaws.com/groups/global/AllUsers"**
`object_acl.Grants.Permission`	Grants.Permission	`get-object-acl`	String	The ACL grantee permission.	Search for S3 object ACLs that grant full control: • **object_acl.Grants.Permission: "FULL_CONTROL"**
`object_acl.Owner.DisplayName`	Owner.DisplayName	`get-object-acl`	String	The ACL owner name.	Search for S3 objects owned by John Doe: • **object_acl.Owner.DisplayName "JohnDoe"**
`object_acl.Owner.ID`	Owner.ID	`get-object-acl`	String	The ACL owner ID.	Search for S3 objects belonging to a particular owner ID: • **object_acl.Owner.ID: "447fba12b05da301df359096f**

S3 Object Data FieldsThat Macie Generates

Macie Field Name	Macie Field Type	Description	Example Search Query
`@timestamp`	Date	The timestamp when the S3 object was last modified.	Search for S3 objects classified by Macie in the last 24 hours: • **@timestamp:[now-1d TO now]**
`content_type`	String	The content and encoding type of the S3 object. **Note** You can locate this value in the **Name** field for a particular content type in the **Content types** section of the Macie	Search for java source code containing hard-coded AWS credentials: • **content_type:"text/x-java-source" AND regex_themes: "aws_access_key"**

Macie Field Name	Macie Field Type	Description	Example Search Query
		console's **Settings** page.	• **content_type:"text/x-java-source" AND regex_themes: "aws_access_key"**
`dlp_risk`	Integer	Through the automatic classification methods, an object monitored by Macie is assigned risk levels based on each content type, file extension, theme, regex, and SVM artifact that is assigned to it. The object's compound (final) risk level (`dlp_risk`) is set to the highest value of its assigned risk levels. **Note** You can find risk levels in the **Settings** page of the Macie console for their respective supported data classifiers.	Search for globally accessible (read or write) objects with the compound (final) risk level of 5 or higher: • **object_acl.Grants.Grantee.URI: "http://acs.amazonaws.com/groups/global/AllUsers" AND dlp_risk>5**
`encoding`	String	The encoding scheme identified when analyzing the S3 object content.	Search for Unicode text documents: • **encoding: "utf-8"**
`filetype_risk`	Integer	The risk level assigned to an S3 object based on its file extension. **Note** You can find risk levels in the **Settings** page of the Macie console for their respective supported data classifiers.	Search for documents with the assigned file extension risk of greater than 6: • **filetype_risk: > 6**

Macie Field Name	Macie Field Type	Description	Example Search Query
`filetypes`	String	The type of the file based on the extension. **Note** You can locate this value in the **Name** and **Description** fields for a particular file type in the **File extensions** section of the Macie console's **Settings** page.	Search for files with an extension of .pdf: • **filetypes: "Adobe PDF (.pdf)"**
`keyword_themes`	String	The themes assigned to an S3 object. You can find supported themes in the Macie console's **Settings** page.	Search for S3 objects containing content related to Social Security: • **keyword_themes: "Social Security Keywords"**
`language_code`	String	The language code found when analyzing the S3 object's content.	Search for S3 objects containing German keywords: • **language_code: "de"**
`last_crawl_time`	Date	The timestamp of when Macie last analyzed an S3 object.	Search for S3 objects analyzed by Macie in the last 24 hours: • **last_crawl_time: [now-1d/d TO now]**
`mimetype_risk`	Integer	The risk level based on an S3 object's content / MIME type.	Search for S3 objects containing MIME types associated with high-risk content: • **mimetype_risk: > 5**
`mimetypes`	String	The MIME type of an S3 object.	Search for plaintext documents containing AWS secret keys: • **mimetypes: "Plain Text (text/plain)" AND themes: aws_secret_key**
`pii_impact`	String	The PII severity impact of an S3 object, assigned by Macie.	Search for S3 objects containing highly valuable personal identifiable information: • **pii_impact: "high"**

Macie Field Name	Macie Field Type	Description	Example Search Query
`pii_types`	String	The specific type of PII found in an S3 object.	Search for S3 objects containing emails: • **pii_types: "email"**
`regex_risk`	Integer	The risk level based on the regex, assigned by Macie, of an S3 object.	Search for S3 objects with a regex-based risk level greater than 5: • **regex_risk: > 5**
`regex_themes`	String	The regex themes of an S3 object.	Search for S3 objects containing RSA private keys • **regex_themes: "RSA Private Key"**
`theme_risk`	String	The risk level based on the themes, assigned by Macie, of an S3 object.	Search for S3 objects with a theme-based risk level higher than 5: • **theme_risk: > 5**
`themes`	String	The combined themes of an S3 object.	Search for S3 objects containing RSA private keys: • **themes: "RSA Private Key"**

Controlling Access to Amazon Macie

The master account users have access to the Macie console, where they can configure Macie and use it to monitor and protect the resources in both master and member accounts. For more information about master and member accounts, see Concepts and Terminology (p. 3) and Integrating Member Accounts and Amazon S3 with Amazon Macie (p. 10).

Granting Administrator Access to Macie

For the master account users to be able to use the Macie console, they must be granted the required permissions. To ensure this, use the following policy document to create and attach an IAM policy to any user identity type that belongs to your master Macie account. This policy grants master account users permissions to use the Macie console in its full capacity.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Resource":"*",
         "Action":[
            "macie:*"
         ]
      },
      {
         "Effect":"Allow",
         "Action":"iam:CreateServiceLinkedRole",
         "Resource":"*",
         "Condition":{
            "StringLike":{
               "iam:AWSServiceName":"macie.amazonaws.com"
            }
         }
      }
   ]
}
```

Granting Read-Only Access to Macie

For a user to view any data in the Macie console, they must be granted the required permissions. To grant read-only access, create a custom policy using the following policy document and attach it to a IAM user, group, or role. This policy grants users permissions to only view information in the Macie console.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "macie:Get*",
                "macie:List*",
```

```
                "macie:Describe*"
            ],
            "Resource": [
                "*"
            ],
            "Effect": "Allow"
        }
    ]
}
```

Note
Currently, there is no AWS managed policy that can grant read-only access to Macie.

Managed (Predefined) Policies for Macie

AWS addresses many common use cases by providing standalone IAM policies that AWS creates and administers. Managed policies grant necessary permissions for common use cases so you can avoid having to investigate what permissions are needed.

The following AWS managed policies, which you can attach to users in your AWS account, are specific to Macie:

- `AmazonMacieFullAccess` – Grants full access to Macie.
- `AmazonMacieSetupRole` – Grants Macie with access to your account.
- `AmazonMacieServiceRole` – Grants Macie read-only access to resource dependencies in your account to enable data analysis.

Disabling Amazon Macie and Deleting Collected Metadata

Use the Macie general settings page in the Macie console to disable Macie.

> **Important**
> Only the master Macie account can disable Macie. For Macie to be disabled in a member account, the master account must disassociate this member account from Macie.

If you disable Macie, it no longer has access to the resources in the master account and all member accounts. You must add member accounts again if you decide to reenable Macie.

If you disable Macie, it stops processing the resources in the master account and all member accounts. After Macie is disabled, the metadata that Macie collected while monitoring the data in your master and member accounts is deleted. Within 90 days from disabling Macie, all of this metadata is expired from the Macie system backups.

> **Important**
> Disabling Macie doesn't prompt the deletion of your other data in your AWS accounts. After Macie is disabled, only the metadata that was collected by Macie while it monitored your accounts is deleted.

1. Navigate to the **Macie general settings** page by choosing the down arrow next to your signed-in name.
2. On the **Macie general settings** page, select the following check boxes:
   - **I understand that if I disable Macie, the service will no longer have access to the resources in the master account and all member accounts. You must add member accounts again if you decide to re-enable Macie.**
   - **I understand that if I disable Macie, the service will stop processing the resources in the master account and all member accounts. All metadata that Macie collected while monitoring the data in these accounts will be deleted.**
3. Choose **Disable Amazon Macie**.

Monitoring Amazon Macie Alerts with Amazon CloudWatch Events

Amazon Macie sends notifications based on CloudWatch Events when any change in the Macie alerts takes place. This includes newly generated alerts and updates to existing alerts. Notifications are sent for all Macie alert types, including predictive alerts and basic alerts, both managed and custom. For more information about alert types, see Amazon Macie Alerts (p. 41).

Macie sends notifications based on CloudWatch Events for the alerts generated in both master and member Macie accounts. However, only the master Macie account has access to the generated events in CloudWatch Events. For more information about master and member accounts, see Concepts and Terminology (p. 3).

The event for Macie in CloudWatch Events has the following format.

Note
In the following example format, the fictional account ID of "111122223333" represents the ID of the master Macie account.

```
{
  "version": "0",
  "id": "CWE-event-id",
  "detail-type": "Macie Alert",
  "source": "aws.macie",
  "account": "111122223333",
  "time": "2017-04-24T22:28:49Z",
  "region": "us-east-1",
  "resources": [
    "arn:aws:macie:us-east-1:111122223333:trigger/trigger_id/alert/alert_id",
    "arn:aws:macie:us-east-1:111122223333:trigger/trigger_id"
  ],
  "detail": {
    "notification-type": "ALERT_CREATED",
    "name": "Scanning bucket policies",
    "tags": [
      "Custom_Alert",
      "Insider"
    ],
    "url": "https://lb00.us-east-1.macie.aws.amazon.com/111122223333/posts/alert_id",
    "alert-arn": "arn:aws:macie:us-east-1:111122223333:trigger/trigger_id/alert/alert_id",
    "risk-score": 8,
    "trigger": {
      "rule-arn": "arn:aws:macie:us-east-1:111122223333:trigger/trigger_id",
      "alert-type": "basic",
      "created-at": "2017-01-02 19:54:00.644000",
      "description": "Alerting on failed enumeration of large number of bucket policies",
      "risk": 8
    },
    "created-at": "2017-04-18T00:21:12.059000",
    "actor": "555566667777:assumed-role:superawesome:aroaidpldc7nsesfnheji",
    "summary": {ALERT_DETAILS_JSON}
  }
}
```

Complete the following procedure to configure your master Macie account to receive events in CloudWatch Events from Macie and pipe those events into an Amazon Simple Queue Service (Amazon SQS) queue. Before completing this procedure, make sure to create the Amazon SQS queue for the events from Macie. For more information, see Tutorial: Creating an Amazon SQS Queue.

1. Open the Amazon CloudWatch console at https://console.aws.amazon.com/cloudwatch/.
2. In the navigation pane, choose **Events** and then choose **Create rule**.
3. Choose **Edit** and enter the following event pattern for the Macie events.

```
{
   "source": [
   "aws.macie"
  ]
}
```

4. In the **Targets** pane, choose **Add target**, select **SQS queue** in the target dropdown, and specify your queue for the events from Macie.

You should now be able to see Macie alerts in your specified queue in the Amazon SQS console.

Document History for Amazon Macie

update-history-change	update-history-description	update-history-date
Add support for service-linked roles for Macie.	Macie can now use the service-linked role named `AWSServiceRoleForAmazonMacie`. It allows Macie to discover, classify, and protect sensitive data in AWS on your behalf. For more information, see Using Service-Linked Roles for Amazon Macie.	June 28, 2018
Added comprehensive tables of fields that can appear in the results of your CloudTrail, S3 bucket properties, and S3 objects data searches. The tables also include search query examples using described fields.	For more information, see Researching AWS CloudTrail Data, Researching S3 Bucket Properties Data, and Researching S3 Objects Data.	May 4, 2018

The following table describes important changes in each release of the *Macie* User Guide.

Earlier updates

Change	Description	Date
Initial publication	Initial publication of the Amazon Macie User Guide.	August 14, 2017