Amazon Macie: User Guide
Copyright © 2020 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon's trademarks and trade dress may not be used in connection with any product or service that is not Amazon's, in any manner that is likely to cause confusion among customers, or in any manner that disparages or discredits Amazon. All other trademarks not owned by Amazon are the property of their respective owners, who may or may not be affiliated with, connected to, or sponsored by Amazon.
# Table of Contents

What Is Amazon Macie? .................................................................................................................. 1
Features of Amazon Macie ............................................................................................................. 1
  Data Discovery and Classification .............................................................................................. 1
  Data Security ............................................................................................................................. 1
Pricing for Macie ................................................................................................................................... 1
Accessing Macie ............................................................................................................................. 2
Concepts and Terminology ............................................................................................................. 3
Setting Up Amazon Macie ............................................................................................................... 5
  Enable Macie ............................................................................................................................... 5
Integrate Amazon S3 with Macie .................................................................................................... 5
Controlling Access to Amazon Macie ............................................................................................ 6
  Granting Administrator Access to Macie ................................................................................. 7
  Granting Read-Only Access to Macie ...................................................................................... 7
Predefined AWS Managed Policies for Macie .............................................................................. 8
  Creating a Handshake Role ....................................................................................................... 8
Service-Linked Roles .................................................................................................................. 10
  Permissions Granted by the Service-Linked Role .................................................................... 10
  Creating a Service-Linked Role for Macie .......................................................................... 11
  Editing a Service-Linked Role for Macie ............................................................................. 11
  Deleting a Service-Linked Role for Macie .......................................................................... 11
Integrating Member Accounts and Amazon S3 ......................................................................... 13
  Integrating Member Accounts with Macie ........................................................................... 13
  Specifying Data for Macie to Monitor ................................................................................... 13
  Encrypted Objects .................................................................................................................. 14
Classifying Data ............................................................................................................................ 15
  Supported Compression and Archive File Formats .............................................................. 15
  Content Type .......................................................................................................................... 16
  File Extension ........................................................................................................................ 23
  Theme .................................................................................................................................... 26
  Regex ...................................................................................................................................... 28
  Personally Identifiable Information ....................................................................................... 30
  Support Vector Machine–Based Classifier ............................................................................. 31
Object Risk Level ....................................................................................................................... 32
Retention Duration for S3 Metadata ............................................................................................ 33
Protecting Data ............................................................................................................................ 34
  AWS CloudTrail Events ........................................................................................................ 34
  AWS CloudTrail Errors ........................................................................................................ 34
Viewing Data and Activity ........................................................................................................... 35
  Dashboard Metrics .................................................................................................................. 35
Dashboard Views ......................................................................................................................... 35
  S3 Objects for Selected Time Range ...................................................................................... 36
  S3 Objects ............................................................................................................................... 36
  S3 Objects by PII ..................................................................................................................... 37
  S3 Public Objects by Buckets ............................................................................................... 38
  S3 Objects by ACL ................................................................................................................ 38
  CloudTrail Events and Associated Users ............................................................................ 39
  CloudTrail Errors and Associated Users ............................................................................ 39
  Activity Location .................................................................................................................... 40
  AWS CloudTrail Events ........................................................................................................ 41
  Activity ISPs .......................................................................................................................... 41
  AWS CloudTrail User Identity Types ................................................................................... 41
Amazon Macie Alerts ................................................................................................................... 42
  Basic and Predictive Macie Alerts ....................................................................................... 42
Alert Categories in Macie ............................................................................................................. 42
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.

Macie is supported in the following AWS Regions:

- US East (N. Virginia) (us-east-1)
- US West (Oregon) (us-west-2)

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

The Macie console is a browser-based interface to access and use Macie. Sign in to your AWS account and open the Macie console using one of the following links:

- https://us-east-1.redirection.macie.aws.amazon.com/
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 in a week.

For more information about alerts, including alert categories and details about the contents of Macie alerts, see Amazon Macie Alerts (p. 42).

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. 49).
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. Before you can get started with Macie, you must enable Macie.

Tasks

- If you don’t have an AWS account, open https://portal.aws.amazon.com/billing/signup and follow the directions.
- Enable Macie (p. 5)
- Integrate Amazon S3 with Macie (p. 5)

Enable Macie

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).

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 the AWS CloudTrail User Guide.

Prerequisites

- The IAM identity (user, role, group) that you use to enable Macie must have the required permissions. To grant the required permissions, attach the AmazonMacieFullAccess managed policy to this identity. For more information, see Predefined AWS Managed Policies for Macie (p. 8).

To enable Amazon Macie

1. Open the Macie console using one of the following links:
2. Choose Get started.
3. (Optional) When you enable Macie, Macie creates a service-linked role. To view the IAM policy for this role, choose View service role permissions. For more information, see Service-Linked Roles for Amazon Macie (p. 10).
4. Choose Enable Macie.

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. 87).

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 is not required. However, we strongly recommend that you integrate with Amazon S3 as part of setting up
Macie. For more information about how Macie classifies your data, see Classifying Data with Amazon Macie (p. 15).

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.

**Prerequisites**

- The IAM identity (user, role, group) that you use to integrate must have the required permissions. To grant the required permissions, attach the `AmazonMacieFullAccess` managed policy to this identity. For more information, see Predefined AWS Managed Policies for Macie (p. 8).

**To integrate with Amazon S3**

1. Log in to AWS with the credentials of the account that is serving as your Macie master account.
2. Open the Macie console and choose **Integrations** from the navigation pane.
3. Choose **S3 Resources** and choose **Select** next to the account (master or member).
4. On the **Integrate S3 resources with Macie** page, choose **Add**. Select up to 250 Amazon S3 resources from the current AWS Region and then choose **Add**.
5. For **Classification of existing objects**, keep the default setting, **Full**. The one-time classification method is applied only once to all of the existing objects in the selected S3 buckets.

Macie displays the following information for each selected bucket:

- **Total objects** – Total number of objects.
- **Processed estimate** – Total size of the data that Macie will classify.
- **Cost estimate** – Cost estimate for classifying all of the objects.

Macie also displays the following totals across all selected buckets:

- **Total size** – Total size of the data.
- **Total number of objects** – Total number of objects.
- **Processed estimate** – Total size of the data that Macie will classify.
- **Total cost estimate** – Cost estimate for classifying all of the objects.

The cost estimate for each bucket is based on its processed estimate value. The total cost estimates are provided only for S3 buckets, not for prefixes. For more information, see Amazon Macie Pricing.

The one-time classification cost estimates are only calculated per S3 bucket, not bucket prefixes. If you select a bucket prefix, the cost estimate for the entire S3 bucket is included in the total cost estimate. If you select multiple prefixes of the same S3 bucket, the cost estimate for the entire S3 bucket is included only once in the total cost estimate.

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

### Controlling Access to Amazon Macie

AWS uses security credentials to identify you and to grant you access to your AWS resources. You can use features of AWS Identity and Access Management (IAM) to allow other users, services, and applications to use your AWS resources fully or in a limited way. You can do this without sharing your security credentials.
By default, IAM users don't have permission to create, view, or modify AWS resources. To allow an IAM user to access resources such as a load balancer, and to perform tasks, you:

1. Create an IAM policy that grants the IAM user permission to use the specific resources and API actions they need.
2. Attach the policy to the IAM user or the group that the IAM user belongs to.

When you attach a policy to a user or group of users, it allows or denies the users permission to perform the specified tasks on the specified resources.

For example, you can use IAM to create users and groups under your AWS account. An IAM user can be a person, a system, or an application. Then you grant permissions to the users and groups to perform specific actions on the specified resources using an IAM policy.

For more information, see the IAM User Guide.

Granting Administrator Access to 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. 13).

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.

```
{
```

7
Predefined AWS Managed Policies for Macie

The managed policies created by AWS grant the required permissions for common use cases. You can attach these policies to IAM users in your AWS account, based on the access to Macie that they require:

- **AmazonMacieFullAccess** – Grants full access to Macie
- **AmazonMacieHandshakeRole** – Grants permission to create the service-linked role for Macie

The following are legacy policies that have been replaced by a service-linked role. For more information, see Legacy Roles for Macie (p. 11).

- **AmazonMacieServiceRole** – Grants Macie read-only access to resource dependencies in your account in order to enable data analysis
- **AmazonMacieSetupRole** – Grants Macie access to your AWS account

Creating a Handshake Role

You can create a role that grants the permissions in the **AmazonMacieHandshakeRole** policy to Macie from the master account as follows.

**To create AWSMacieServiceCustomerHandshakeRole using the IAM console**

1. Open the IAM console at https://console.aws.amazon.com/iam/.
2. In the navigation pane, choose **Roles**.
3. Choose **Create role** and do the following:
   a. For **Select type of trusted entity**, choose **AWS service**.
   b. For **Choose a use case**, select **EC2**.
   c. Choose **Next: Permissions**.
4. On the **Attach permissions policies** page, select the checkbox for the **AmazonMacieHandshakeRole** policy and choose **Next: Tags**.
5. (Optional) Add tags to your role and then choose **Next: Review**.
6. On the **Review** page, do the following:
   a. For **Role name**, enter **AWSMacieServiceCustomerHandshakeRole**.
   b. For **Role description**, enter the following: Allows the master account to create service-linked roles in the member accounts.
   c. Choose **Create role**.
7. Edit the trust policy as follows:

   a. Select AWSMacieServiceCustomerHandshakeRole, which you just created.
   b. On the Trust relationships tab, choose Edit trust relationship.
   c. Enter the following trust policy:

   ```json
   {
   "Version": "2012-10-17",
   "Statement": [
      {
      "Effect": "Allow",
      "Principal": {
      "Service": "macie.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
      "StringEquals": {
      "sts:ExternalId": "master-account-id"
      }
      }
      }
   ]
   }
   ```
   d. Choose Update Trust Policy.

**To create AWSMacieServiceCustomerHandshakeRole using the AWS CLI**

1. Create the following trust policy and save it in a text file named `macie-handshake-trust-policy.json`.

   ```json
   {
   "Version": "2012-10-17",
   "Statement": [
      {
      "Effect": "Allow",
      "Principal": {
      "Service": "macie.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
      "StringEquals": {
      "sts:ExternalId": "master-account-id"
      }
      }
      }
   ]
   }
   ```

2. Create the role and specify the trust policy that you created in the previous step using the `create-role` command.

   ```bash
   aws iam create-role --role-name AWSMacieServiceCustomerHandshakeRole --assume-role-policy-document file://macie-handshake-trust-policy.json
   ```

3. Attach the AmazonMacieHandshakeRole policy to the role using the `attach-role-policy` command.

   ```bash
   aws iam attach-role-policy --role-name AWSMacieServiceCustomerHandshakeRole --policy-arn arn:aws:iam::aws:policy/service-role/AmazonMacieHandshakeRole
   ```
Service-Linked Roles for Amazon Macie

Amazon Macie uses AWS Identity and Access Management (IAM) service-linked roles to call other AWS services on your behalf. Service-linked roles provide a secure way to delegate permissions to Macie because only Macie can assume the service-linked role.

Permissions Granted by the Service-Linked Role

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 role trusts the macie.amazonaws.com service to assume it.

The role is configured with the following AWS managed policy.

```json
{
   "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",
         "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"
         ]
      }
   ]
}
```
Creating a Service-Linked Role for Macie

You don't need to manually create the \texttt{AWSServiceRoleForAmazonMacie} role. Macie creates this role on your behalf as follows:

- **Master account** — The \texttt{AWSServiceRoleForAmazonMacie} role is created for you when you enable Macie for the first time.
- **Member accounts** — The \texttt{AWSServiceRoleForAmazonMacie} role is created for you when the master account associates the member account with Macie. The service-linked role that is created for the master Macie account doesn't apply to the member Macie accounts.

For Macie to create the service-linked role on your behalf, you must have the required permissions. To grant the required permissions to an IAM entity, such as a user, group, or role, attach the \texttt{AmazonMacieFullAccess} policy. For more information, see Predefined AWS Managed Policies for Macie (p. 8) and Service-Linked Role Permissions in the IAM User Guide.

You can also create the \texttt{AWSServiceRoleForAmazonMacie} role manually. For more information, see Creating a Service-Linked Role in the IAM User Guide.

Legacy Roles for Macie

If you used Macie 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 \texttt{AmazonMacieServiceRole} and \texttt{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 service-linked role replaces these previously created IAM roles (in master and member accounts). The previously created roles were not deleted, but they're no longer used to grant Macie permission to call other services on your behalf. You can use the IAM console to delete the previously created roles.

Editing a Service-Linked Role for Macie

After you create a service-linked role, you can't change the name of 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 Amazon Macie, we recommend that you delete the \texttt{AWSServiceRoleForAmazonMacie} role.

For a master account, you can delete the Macie service-linked role only after disabling Macie. This ensures that you can't inadvertently remove permissions to access Macie resources. For member accounts, the master account must first disassociate them from Macie. For more information, see Disabling Amazon Macie and Deleting Collected Metadata (p. 87).

When you disable Macie, the \texttt{AWSServiceRoleForAmazonMacie} role is not deleted. If you enable Macie again, it uses the existing \texttt{AWSServiceRoleForAmazonMacie} role.
You can use the IAM console, the IAM CLI, or the IAM API to delete the `AWSServiceRoleForAmazonMacie` 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 integrate member accounts with Macie and integrate Amazon S3 with Macie. For more information about the master and member accounts, see Concepts and Terminology (p. 3).

Contents

• Integrating Member Accounts with Macie (p. 13)
• Specifying Data for Macie to Monitor (p. 13)
• Encrypted Objects (p. 14)

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.

Prerequisites

• Create a role that grants the member account the permissions required to create the AWSServiceRoleForAmazonMacie service-linked role. For more information, see Creating a Handshake Role (p. 8).

To integrate member accounts with Macie

1. Log in to AWS with the credentials of the AWS account that is serving as your Macie master account.
2. Open the Macie console and choose Integrations from the navigation pane.
3. Choose Accounts and choose the plus icon (+) next to Member AWS accounts.
4. When prompted, enter one or more account IDs, separated by commas. Choose Add accounts.
5. (Optional) Verify that Macie created the AWSServiceRoleForAmazonMacie role in each member account that you integrated. For more information, see Creating a Service-Linked Role for Macie (p. 11).

Specifying Data for Macie to Monitor

You can specify the S3 buckets and prefixes that contain the data for Macie to monitor.

Prerequisites

• The IAM identity (user, role, group) that you use to integrate must have the required permissions. To grant the required permissions, attach the AmazonMacieFullAccess managed policy to this identity. For more information, see Predefined AWS Managed Policies for Macie (p. 8).

To update your integration with Amazon S3

1. Log in to AWS with the credentials of the account that is serving as your Macie master account.
2. Open the Macie console and choose **Integrations** from the navigation pane.
3. Choose **S3 Resources** and choose **Select** next to the account (master or member).
4. On the **Integrate S3 resources with Macie** page, choose **Edit** to edit the buckets/prefixes that are already integrated or **Add** to integrate new buckets/prefixes.
5. For **Classification of existing objects**, keep the default setting, **Full**. The one-time classification method is applied only once to all of the existing objects in the selected S3 buckets.

Macie displays the following information for each selected bucket:

- **Total objects** – Total number of objects.
- **Processed estimate** – Total size of the data that Macie will classify.
- **Cost estimate** – Cost estimate for classifying all of the objects.

Macie also displays the following totals across all selected buckets:

- **Total size** – Total size of the data.
- **Total number of objects** – Total number of objects.
- **Processed estimate** – Total size of the data that Macie will classify.
- **Total cost estimate** – Cost estimate for classifying all of the objects.

The cost estimate for each bucket is based on its processed estimate value. The total cost estimates are provided only for S3 buckets, not for prefixes. For more information, see [Amazon Macie Pricing](https://aws.amazon.com/macie/pricing/).

The one-time classification cost estimates are only calculated per S3 bucket, not bucket prefixes. If you select a bucket prefix, the cost estimate for the entire S3 bucket is included in the total cost estimate. If you select multiple prefixes of the same S3 bucket, the cost estimate for the entire S3 bucket is included only once in the total cost estimate.

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 `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.

After 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, see Specifying Data for Macie to Monitor (p. 13).

Limits

- Macie has a default limit on the amount of data that it can classify in an account. After 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.
- If you specify S3 buckets that include files of a format that isn't supported in Macie, Macie doesn't classify them.
- Macie's content classification engine processes up to the first 20 MB of an S3 object.

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.

Contents

- Supported Compression and Archive File Formats (p. 15)
- Content Type (p. 16)
- File Extension (p. 23)
- Theme (p. 26)
- Regex (p. 28)
- Personally Identifiable Information (p. 30)
- Support Vector Machine–Based Classifier (p. 31)
- Object Risk Level (p. 32)
- Retention Duration for S3 Metadata (p. 33)

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
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, with 10 being the highest risk and 1 being the lowest.

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.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>application/cap</td>
<td>WireShark or Tcpdump Packet Capture</td>
</tr>
<tr>
<td>application/epub+zip</td>
<td>application/epub</td>
</tr>
<tr>
<td>application/illustrator</td>
<td>Adobe Illustrator</td>
</tr>
<tr>
<td>application/java</td>
<td>Binary (Java)</td>
</tr>
<tr>
<td>application/java-archive</td>
<td>application/java-archive</td>
</tr>
<tr>
<td>application/java-serialized-object</td>
<td>application/java-serialized-object</td>
</tr>
<tr>
<td>application/java-vm</td>
<td>application/java-vm</td>
</tr>
<tr>
<td>application/javascript</td>
<td>application/javascript</td>
</tr>
<tr>
<td>application/json</td>
<td>JSON</td>
</tr>
<tr>
<td>Content Type</td>
<td>Description</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>application/msaccess</td>
<td>application/msaccess</td>
</tr>
<tr>
<td>application/msexcel</td>
<td>Microsoft Excel</td>
</tr>
<tr>
<td>application/msonenote</td>
<td>application/msonenote</td>
</tr>
<tr>
<td>application/mspowerpoint</td>
<td>Microsoft PowerPoint</td>
</tr>
<tr>
<td>application/msword</td>
<td>Microsoft Word</td>
</tr>
<tr>
<td>application/octet-stream</td>
<td>application/octet-stream</td>
</tr>
<tr>
<td>application/octet-stream+fon</td>
<td>application/octet-stream+fon</td>
</tr>
<tr>
<td>application/ogg</td>
<td>application/ogg</td>
</tr>
<tr>
<td>application/onenote</td>
<td>application/onenote</td>
</tr>
<tr>
<td>application/pdf</td>
<td>Adobe PDF</td>
</tr>
<tr>
<td>application/pgp</td>
<td>application/pgp</td>
</tr>
<tr>
<td>application/pgp-encrypted</td>
<td>application/pgp-encrypted</td>
</tr>
<tr>
<td>application/pgp-signature</td>
<td>PGP signature</td>
</tr>
<tr>
<td>application/postscript</td>
<td>Adobe Postscript</td>
</tr>
<tr>
<td>application/rar</td>
<td>RAR compressed archive</td>
</tr>
<tr>
<td>application/rdf+xml</td>
<td>application/rdf+xml</td>
</tr>
<tr>
<td>application/rss+xml</td>
<td>application/rss+xml</td>
</tr>
<tr>
<td>application/rtf</td>
<td>application/rtf</td>
</tr>
<tr>
<td>application/tar</td>
<td>TAR archive</td>
</tr>
<tr>
<td>application/unknown</td>
<td>application/unknown</td>
</tr>
<tr>
<td>application/vnd.3gpp.pic-bw-small</td>
<td>application/vnd.3gpp.pic-bw-small</td>
</tr>
<tr>
<td>application/vnd.android.package-archive</td>
<td>Android Package</td>
</tr>
<tr>
<td>application/vnd.audiograph</td>
<td>application/vnd.audiograph</td>
</tr>
<tr>
<td>application/vnd.balsamiq.bmpr</td>
<td>Balsamiq Mockup</td>
</tr>
<tr>
<td>application/vnd.curl.car</td>
<td>application/vnd.curl.car</td>
</tr>
<tr>
<td>application/vnd.dvb.ait</td>
<td>application/vnd.dvb.ait</td>
</tr>
<tr>
<td>application/vnd.google-apps.document</td>
<td>Google Apps Document</td>
</tr>
<tr>
<td>application/vnd.google-apps.drawing</td>
<td>application/vnd.google-apps.drawing</td>
</tr>
<tr>
<td>application/vnd.google-apps.form</td>
<td>Google Apps Form</td>
</tr>
<tr>
<td>application/vnd.google-apps.map</td>
<td>Google Apps Map</td>
</tr>
<tr>
<td>application/vnd.google-apps.presentation</td>
<td>Google Apps Presentation</td>
</tr>
<tr>
<td>Content Type</td>
<td>Description</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>application/vnd.ms-word.document.macroEnabled.12</td>
<td>Microsoft Word - Macro enabled</td>
</tr>
<tr>
<td>application/vnd.ms-xpsdocument</td>
<td></td>
</tr>
<tr>
<td>application/vnd.oasis.opendocument.chart</td>
<td></td>
</tr>
<tr>
<td>application/vnd.oasis.opendocument.graphics</td>
<td></td>
</tr>
<tr>
<td>application/vnd.oasis.opendocument.presentation</td>
<td>Presentation</td>
</tr>
<tr>
<td>application/vnd.oasis.opendocument.spreadsheet</td>
<td>Spreadsheet</td>
</tr>
<tr>
<td>application/vnd.oasis.opendocument.text</td>
<td>Open Document Text</td>
</tr>
<tr>
<td>application/vnd.openxmlformats-officedocument.presentationml.presentation</td>
<td>Microsoft PowerPoint</td>
</tr>
<tr>
<td>application/vnd.openxmlformats-officedocument.presentationml.slide</td>
<td>Microsoft Powerpoint</td>
</tr>
<tr>
<td>application/vnd.openxmlformats-officedocument.presentationml.slideshow</td>
<td>Microsoft Powerpoint</td>
</tr>
<tr>
<td>application/vnd.openxmlformats-officedocument.presentationml.template</td>
<td>Microsoft Excel</td>
</tr>
<tr>
<td>application/vnd.openxmlformats-officedocument.spreadsheetml.sheet</td>
<td>application/vnd.openxmlformats-officedocument.spreadsheetml.template</td>
</tr>
<tr>
<td>application/vnd.openxmlformats-officedocument.spreadsheetml.template</td>
<td>Microsoft Excel</td>
</tr>
<tr>
<td>application/vnd.openxmlformats-officedocument.wordprocessingml.document</td>
<td>Microsoft Word</td>
</tr>
<tr>
<td>application/vnd.openxmlformats-officedocument.wordprocessingml.template</td>
<td>application/vnd.openxmlformats-officedocument.wordprocessingml.template</td>
</tr>
<tr>
<td>application/vnd.palm</td>
<td>application/vnd.palm</td>
</tr>
<tr>
<td>application/vnd.symbian.install</td>
<td>application/vnd.symbian.install</td>
</tr>
<tr>
<td>application/x-7z-compressed</td>
<td>7zip compressed archive</td>
</tr>
<tr>
<td>application/x-adobeamdetect</td>
<td>Adobe Application Manager</td>
</tr>
<tr>
<td>application/x-adobeexmandetect</td>
<td>application/x-adobeexmandetect</td>
</tr>
<tr>
<td>application/x-apple-diskimage</td>
<td>Apple disk image</td>
</tr>
<tr>
<td>application/x-bittorrent</td>
<td>application/x-bittorrent</td>
</tr>
<tr>
<td>application/x-bzip2</td>
<td>application/x-bzip2</td>
</tr>
<tr>
<td>Content Type</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>application/x-cab</td>
<td>application/x-cab</td>
</tr>
<tr>
<td>application/x-cfs-compressed</td>
<td>application/x-cfs-compressed</td>
</tr>
<tr>
<td>application/x-coredump</td>
<td>application/x-coredump</td>
</tr>
<tr>
<td>application/x-couponprinterplugin</td>
<td>application/x-couponprinterplugin</td>
</tr>
<tr>
<td>application/x-dbm</td>
<td>application/x-dbm</td>
</tr>
<tr>
<td>application/x-doseexec</td>
<td>Executable</td>
</tr>
<tr>
<td>application/x-dvi</td>
<td>application/x-dvi</td>
</tr>
<tr>
<td>application/x-executable</td>
<td>Executable</td>
</tr>
<tr>
<td>application/x-fla</td>
<td>application/x-fla</td>
</tr>
<tr>
<td>application/x-font</td>
<td>application/x-font</td>
</tr>
<tr>
<td>application/x-font-otf</td>
<td>application/x-font-otf</td>
</tr>
<tr>
<td>application/x-font-ttf</td>
<td>application/x-font-ttf</td>
</tr>
<tr>
<td>application/x-font-type1</td>
<td>application/x-font-type1</td>
</tr>
<tr>
<td>application/x-font-woff</td>
<td>application/x-font-woff</td>
</tr>
<tr>
<td>application/x-freemind</td>
<td>application/x-freemind</td>
</tr>
<tr>
<td>application/x-gtar</td>
<td>GNU tar compressed archive</td>
</tr>
<tr>
<td>application/x-gzip</td>
<td>GNU Zip compressed archive</td>
</tr>
<tr>
<td>application/x-iso9660-image</td>
<td>application/x-iso9660-image</td>
</tr>
<tr>
<td>application/x-iwork-keynote-sffkey</td>
<td>application/x-iwork-keynote-sffkey</td>
</tr>
<tr>
<td>application/x-iwork-numbers-sffnumbers</td>
<td>application/x-iwork-numbers-sffnumbers</td>
</tr>
<tr>
<td>application/x-iwork-pages-sffpages</td>
<td>application/x-iwork-pages-sffpages</td>
</tr>
<tr>
<td>application/x-javascript</td>
<td>application/x-javascript</td>
</tr>
<tr>
<td>application/x-maker</td>
<td>application/x-maker</td>
</tr>
<tr>
<td>application/x-mobipocket-ebook</td>
<td>application/x-mobipocket-ebook</td>
</tr>
<tr>
<td>application/x-ms-shortcut</td>
<td>application/x-ms-shortcut</td>
</tr>
<tr>
<td>application/x-ms-wmz</td>
<td>application/x-ms-wmz</td>
</tr>
<tr>
<td>application/x-msdos-program</td>
<td>Microsoft Windows Application</td>
</tr>
<tr>
<td>application/x-msi</td>
<td>application/x-msi</td>
</tr>
<tr>
<td>application/x-msmetafile</td>
<td>application/x-msmetafile</td>
</tr>
<tr>
<td>application/x-mspublisher</td>
<td>application/x-mspublisher</td>
</tr>
<tr>
<td>application/x-nawk</td>
<td>application/x-nawk</td>
</tr>
<tr>
<td>application/x-ns-proxy-autoconfig</td>
<td>application/x-ns-proxy-autoconfig</td>
</tr>
<tr>
<td>application/x-object</td>
<td>application/x-object</td>
</tr>
<tr>
<td>Content Type</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>application/x-perl</td>
<td>Perl Source Code</td>
</tr>
<tr>
<td>application/x-pkcs12</td>
<td>PKI Certificate</td>
</tr>
<tr>
<td>application/x-pkcs7-crl</td>
<td>PKI Files</td>
</tr>
<tr>
<td>application/x-python-code</td>
<td>Source Code (python)</td>
</tr>
<tr>
<td>application/x-rar-compressed</td>
<td>RAR compressed archive</td>
</tr>
<tr>
<td>application/x-redhat-package-manager</td>
<td>application/x-redhat-package-manager</td>
</tr>
<tr>
<td>application/x-sas</td>
<td>Statistical Analysis</td>
</tr>
<tr>
<td>application/x-sharedlib</td>
<td>application/x-sharedlib</td>
</tr>
<tr>
<td>application/x-shellscript</td>
<td>Shell Script</td>
</tr>
<tr>
<td>application/x-shockwave-flash</td>
<td>application/x-shockwave-flash</td>
</tr>
<tr>
<td>application/x-silverlight-app</td>
<td>application/x-silverlight-app</td>
</tr>
<tr>
<td>application/x-stuffit</td>
<td>Stuffit compressed archive</td>
</tr>
<tr>
<td>application/x-subrip</td>
<td>application/x-subrip</td>
</tr>
<tr>
<td>application/x-tar</td>
<td>TAR archive</td>
</tr>
<tr>
<td>application/x-tex-tfm</td>
<td>Apache Font</td>
</tr>
<tr>
<td>application/x-texinfo</td>
<td>application/x-texinfo</td>
</tr>
<tr>
<td>application/x-troff-man</td>
<td>application/x-troff-man</td>
</tr>
<tr>
<td>application/x-wais-source</td>
<td>application/x-wais-source</td>
</tr>
<tr>
<td>application/x-x509-ca-cert</td>
<td>application/x-x509-ca-cert</td>
</tr>
<tr>
<td>application/x-xcf</td>
<td>application/x-xcf</td>
</tr>
<tr>
<td>application/x-xfig</td>
<td>application/x-xfig</td>
</tr>
<tr>
<td>application/x-xpinstall</td>
<td>application/x-xpinstall</td>
</tr>
<tr>
<td>application/x-xminda</td>
<td>Zip compressed archive</td>
</tr>
<tr>
<td>application/xhtml+xml</td>
<td>application/xhtml+xml</td>
</tr>
<tr>
<td>application/xml</td>
<td>application/xml</td>
</tr>
<tr>
<td>application/xv+xml</td>
<td>application/xv+xml</td>
</tr>
<tr>
<td>application/zip</td>
<td>Zip compressed archive</td>
</tr>
<tr>
<td>binary/octet-stream</td>
<td>binary/octet-stream</td>
</tr>
<tr>
<td>chemical/x-cache</td>
<td>chemical/x-cache</td>
</tr>
<tr>
<td>chemical/x-cerius</td>
<td>chemical/x-cerius</td>
</tr>
<tr>
<td>chemical/x-gamess-input</td>
<td>chemical/x-gamess-input</td>
</tr>
<tr>
<td>chemical/x-genbank</td>
<td>chemical/x-genbank</td>
</tr>
<tr>
<td>Content Type</td>
<td>Description</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>chemical/x-mdl-sdfile</td>
<td>chemical/x-mdl-sdfile</td>
</tr>
<tr>
<td>chemical/x-pdb</td>
<td>Protein Databank chemical/x-pdb</td>
</tr>
<tr>
<td>chemical/x-rosdal</td>
<td>chemical/x-rosdal</td>
</tr>
<tr>
<td>message/rfc822</td>
<td>message/rfc822</td>
</tr>
<tr>
<td>text/cache-manifest</td>
<td>text/cache-manifest</td>
</tr>
<tr>
<td>text/calendar</td>
<td>text/calendar</td>
</tr>
<tr>
<td>text/css</td>
<td>text/css</td>
</tr>
<tr>
<td>text/csv</td>
<td>Comma Separated Values</td>
</tr>
<tr>
<td>text/html</td>
<td>text/html</td>
</tr>
<tr>
<td>text/json</td>
<td>JavaScript Object Notation</td>
</tr>
<tr>
<td>text/plain</td>
<td>Plain Text</td>
</tr>
<tr>
<td>text/rtf</td>
<td>text/rtf</td>
</tr>
<tr>
<td>text/tab-separated-values</td>
<td>Tab separated values</td>
</tr>
<tr>
<td>text/texmacs</td>
<td>text/texmacs</td>
</tr>
<tr>
<td>text/vnd.graphviz</td>
<td>text/vnd.graphviz</td>
</tr>
<tr>
<td>text/x-asn</td>
<td>Source Code (Assembly)</td>
</tr>
<tr>
<td>text/x-bibtex</td>
<td>text/x-bibtex</td>
</tr>
<tr>
<td>text/x-c</td>
<td>Source Code (c)</td>
</tr>
<tr>
<td>text/x-c++hdr</td>
<td>Source Code (C++ headers)</td>
</tr>
<tr>
<td>text/x-c++src</td>
<td>Source Code (c++)</td>
</tr>
<tr>
<td>text/x-chdr</td>
<td>Source Code (C headers)</td>
</tr>
<tr>
<td>text/x-component</td>
<td>text/x-component</td>
</tr>
<tr>
<td>text/x-csh</td>
<td>Source Code (C shell)</td>
</tr>
<tr>
<td>text/x-csharp</td>
<td>Source Code (C#)</td>
</tr>
<tr>
<td>text/x-csrc</td>
<td>Source Code (C)</td>
</tr>
<tr>
<td>text/x-diff</td>
<td>text/x-diff</td>
</tr>
<tr>
<td>text/x-dsrec</td>
<td>text/x-dsrec</td>
</tr>
<tr>
<td>text/x-java</td>
<td>Source Code (Java)</td>
</tr>
<tr>
<td>text/x-java-source</td>
<td>Source Code (Java)</td>
</tr>
<tr>
<td>text/x-markdown</td>
<td>text/x-markdown</td>
</tr>
<tr>
<td>text/x-nfo</td>
<td>text/x-nfo</td>
</tr>
<tr>
<td>text/x-objcsrc</td>
<td>Source Code (Objective-C)</td>
</tr>
<tr>
<td>text/x-pascal</td>
<td>Source Code (Pascal)</td>
</tr>
</tbody>
</table>
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, with 10 being the highest risk and 1 being the lowest.

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.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>7z</td>
<td>7-Zip compressed file</td>
</tr>
<tr>
<td>abc</td>
<td>SolidWorks CAD</td>
</tr>
<tr>
<td>accdb</td>
<td>Microsoft Access database</td>
</tr>
<tr>
<td>apk</td>
<td>Application installable on Android</td>
</tr>
<tr>
<td>bat</td>
<td>Batch file</td>
</tr>
<tr>
<td>text/x-perl</td>
<td>Source Code (Perl)</td>
</tr>
<tr>
<td>text/x-python</td>
<td>Source Code (Python)</td>
</tr>
<tr>
<td>text/x-sfv</td>
<td>text/x-sfv</td>
</tr>
<tr>
<td>text/x-sh</td>
<td>Source Code (x-sh)</td>
</tr>
<tr>
<td>text/x-sql</td>
<td>Source Code (SQL)</td>
</tr>
<tr>
<td>text/x-tex</td>
<td>text/x-tex</td>
</tr>
<tr>
<td>text/x-url</td>
<td>text/x-url</td>
</tr>
<tr>
<td>text/x-vcard</td>
<td>text/x-vcard</td>
</tr>
<tr>
<td>text/xml</td>
<td>XML Text</td>
</tr>
<tr>
<td>File Extension</td>
<td>Description</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td>bin</td>
<td>Compressed archive. Readable by Java. Extractable by 7-zip</td>
</tr>
<tr>
<td>bz2</td>
<td>Bzip2 compressed archive</td>
</tr>
<tr>
<td>bzip2</td>
<td>Bzip2 compressed archive</td>
</tr>
<tr>
<td>c</td>
<td>C source code</td>
</tr>
<tr>
<td>c#</td>
<td>C# source code</td>
</tr>
<tr>
<td>cab</td>
<td>Microsoft cabinet. Extractable via ZIP</td>
</tr>
<tr>
<td>cc</td>
<td>C++ source code</td>
</tr>
<tr>
<td>cer</td>
<td>PKI certificate</td>
</tr>
<tr>
<td>cpp</td>
<td>C++ source code</td>
</tr>
<tr>
<td>csv</td>
<td>Comma Separated Values</td>
</tr>
<tr>
<td>cxx</td>
<td>C++ source code</td>
</tr>
<tr>
<td>dbf</td>
<td>dBase database</td>
</tr>
<tr>
<td>dbx</td>
<td>Microsoft Outlook Express</td>
</tr>
<tr>
<td>deb</td>
<td>Debian Linux install package</td>
</tr>
<tr>
<td>dmg</td>
<td>Apple OS X Application Installer</td>
</tr>
<tr>
<td>doc</td>
<td>Microsoft Word</td>
</tr>
<tr>
<td>docx</td>
<td>Microsoft Word</td>
</tr>
<tr>
<td>dot</td>
<td>Microsoft Word</td>
</tr>
<tr>
<td>dotx</td>
<td>Microsoft Word</td>
</tr>
<tr>
<td>dwg</td>
<td>AutoDesk CAD</td>
</tr>
<tr>
<td>dxf</td>
<td>AutoCAD</td>
</tr>
<tr>
<td>eml</td>
<td>MIME email</td>
</tr>
<tr>
<td>emlx</td>
<td>Apple Mail email message</td>
</tr>
<tr>
<td>exe</td>
<td>Microsoft Windows PE Executable</td>
</tr>
<tr>
<td>gpg</td>
<td>PGP certificate</td>
</tr>
<tr>
<td>gz</td>
<td>GNU Zip compressed archive</td>
</tr>
<tr>
<td>gzip</td>
<td>GNU Zip compressed archive</td>
</tr>
<tr>
<td>html</td>
<td>Hyper Text Markup Language</td>
</tr>
<tr>
<td>iwa</td>
<td>Apple iWork document archive file</td>
</tr>
<tr>
<td>jar</td>
<td>Java source code archive</td>
</tr>
<tr>
<td>java</td>
<td>Java source code</td>
</tr>
<tr>
<td>json</td>
<td>Java Script Object Notation Values (JSON)</td>
</tr>
<tr>
<td>key</td>
<td>Apple Keynote Presentation</td>
</tr>
<tr>
<td>File Extension</td>
<td>Description</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td>keynote</td>
<td>Apple Keynote Presentation</td>
</tr>
<tr>
<td>lua</td>
<td>Lua source code</td>
</tr>
<tr>
<td>mdb</td>
<td>Microsoft Access database</td>
</tr>
<tr>
<td>msg</td>
<td>Microsoft Outlook Message</td>
</tr>
<tr>
<td>msi</td>
<td>Microsoft Windows Application Installer</td>
</tr>
<tr>
<td>odp</td>
<td>OpenOffice.org OpenDocument presentation file</td>
</tr>
<tr>
<td>oos</td>
<td>OpenOffice.org spreadsheet file</td>
</tr>
<tr>
<td>p12</td>
<td>PKI certificate</td>
</tr>
<tr>
<td>pages</td>
<td>Apple Pages</td>
</tr>
<tr>
<td>pdf</td>
<td>Adobe PDF</td>
</tr>
<tr>
<td>perl</td>
<td>Perl source code</td>
</tr>
<tr>
<td>pgp</td>
<td>PGP certificate</td>
</tr>
<tr>
<td>pl</td>
<td>Perl source code</td>
</tr>
<tr>
<td>pot</td>
<td>Microsoft PowerPoint</td>
</tr>
<tr>
<td>pps</td>
<td>Microsoft PowerPoint</td>
</tr>
<tr>
<td>ppt</td>
<td>Microsoft PowerPoint</td>
</tr>
<tr>
<td>pptx</td>
<td>Microsoft PowerPoint</td>
</tr>
<tr>
<td>pst</td>
<td>Microsoft Outlook</td>
</tr>
<tr>
<td>py</td>
<td>Python source code</td>
</tr>
<tr>
<td>rar</td>
<td>RAR archive. Extractable by 7-zip</td>
</tr>
<tr>
<td>rtf</td>
<td>Rich Text Format</td>
</tr>
<tr>
<td>sdp</td>
<td>OpenOffice.org presentation file</td>
</tr>
<tr>
<td>sdw</td>
<td>OpenOffice.org text document file</td>
</tr>
<tr>
<td>sldasm</td>
<td>SolidWorks CAD</td>
</tr>
<tr>
<td>slddrw</td>
<td>SolidWorks CAD</td>
</tr>
<tr>
<td>sldprt</td>
<td>SolidWorks CAD</td>
</tr>
<tr>
<td>sql</td>
<td>Structured Query Language</td>
</tr>
<tr>
<td>sxi</td>
<td>OpenOffice.org presentation file</td>
</tr>
<tr>
<td>sxw</td>
<td>OpenOffice.org Writer document file</td>
</tr>
<tr>
<td>tar.gz</td>
<td>GNU Zip compressed archive</td>
</tr>
<tr>
<td>tsv</td>
<td>Tab Separated Values</td>
</tr>
<tr>
<td>txt</td>
<td>Text Document</td>
</tr>
<tr>
<td>vdx</td>
<td>Microsoft Visio</td>
</tr>
</tbody>
</table>
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, with 10 being the highest risk and 1 being the lowest.

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.

<table>
<thead>
<tr>
<th>Theme title</th>
<th>Minimum keyword combinations</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Express Credit Card Keywords</td>
<td>1</td>
</tr>
<tr>
<td>Attorney Client Privileged</td>
<td>2</td>
</tr>
<tr>
<td>Audit Keywords</td>
<td>3</td>
</tr>
<tr>
<td>Theme</td>
<td>Count</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Banking Keywords</td>
<td>1</td>
</tr>
<tr>
<td>Big Data Frameworks</td>
<td>2</td>
</tr>
<tr>
<td>Cisco Analysis Keywords</td>
<td>1</td>
</tr>
<tr>
<td>Confidential Markings</td>
<td>2</td>
</tr>
<tr>
<td>Corporate Growth Keywords</td>
<td>3</td>
</tr>
<tr>
<td>Corporate Project Plan</td>
<td>3</td>
</tr>
<tr>
<td>Corporate Proposals</td>
<td>3</td>
</tr>
<tr>
<td>Credit Card Keywords</td>
<td>1</td>
</tr>
<tr>
<td>Encrypted Data Keywords</td>
<td>1</td>
</tr>
<tr>
<td>Financial Keywords</td>
<td>1</td>
</tr>
<tr>
<td>Hacker Keywords</td>
<td>2</td>
</tr>
<tr>
<td>Limit Distribution Markings</td>
<td>3</td>
</tr>
<tr>
<td>Mastercard Credit Card Keywords</td>
<td>1</td>
</tr>
<tr>
<td>Metasploit Framework Keywords</td>
<td>1</td>
</tr>
<tr>
<td>NMAP OS Fingerprinting</td>
<td>1</td>
</tr>
<tr>
<td>Network Scanner Keywords</td>
<td>1</td>
</tr>
<tr>
<td>Network Service Fingerprinting Keywords</td>
<td>1</td>
</tr>
<tr>
<td>Network Traffic Analysis Keywords</td>
<td>1</td>
</tr>
<tr>
<td>OS Backdoor Keywords</td>
<td>1</td>
</tr>
<tr>
<td>Offline Attacks Keywords</td>
<td>1</td>
</tr>
<tr>
<td>Online Attacks Keywords</td>
<td>1</td>
</tr>
<tr>
<td>Oracle DB Analysis Keywords</td>
<td>1</td>
</tr>
<tr>
<td>Password Keywords</td>
<td>2</td>
</tr>
<tr>
<td>Project Tracking Keywords</td>
<td>2</td>
</tr>
<tr>
<td>Proprietary Markings</td>
<td>2</td>
</tr>
<tr>
<td>Real-Time Processing Frameworks</td>
<td>2</td>
</tr>
<tr>
<td>Restricted Markings</td>
<td>2</td>
</tr>
<tr>
<td>SSL Forensic Analysis Keywords</td>
<td>1</td>
</tr>
<tr>
<td>Secret Markings</td>
<td>3</td>
</tr>
<tr>
<td>Sensitive Markings</td>
<td>3</td>
</tr>
<tr>
<td>Social Security Keywords</td>
<td>2</td>
</tr>
<tr>
<td>Stock Keywords</td>
<td>3</td>
</tr>
<tr>
<td>Taxpayer EIN Keywords</td>
<td>2</td>
</tr>
</tbody>
</table>
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, with 10 being the highest risk and 1 being the lowest.

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.

<table>
<thead>
<tr>
<th>Name</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arista network configuration</td>
<td>Regex</td>
</tr>
<tr>
<td>BBVA Compass Routing Number - California</td>
<td>Regex</td>
</tr>
<tr>
<td>Bank of America Routing Numbers - California</td>
<td>Regex</td>
</tr>
<tr>
<td>Box Links</td>
<td>Regex</td>
</tr>
<tr>
<td>CVE Number</td>
<td>Regex</td>
</tr>
<tr>
<td>California Drivers License</td>
<td>Regex</td>
</tr>
<tr>
<td>Chase Routing Numbers - California</td>
<td>Regex</td>
</tr>
<tr>
<td>Cisco Router Config</td>
<td>Regex</td>
</tr>
</tbody>
</table>
Citibank Routing Numbers - California
DSA Private Key
Dropbox Links
EC Private Key
Encrypted DSA Private Key
Encrypted EC Private Key
Encrypted Private Key
Encrypted PuTTY SSH DSA Key
Encrypted PuTTY SSH RSA Key
Encrypted RSA Private Key
Google Application Identifier
HIPAA PHI National Drug Code
Huawei config file
Individual Taxpayer Identification Numbers (ITIN)
John the Ripper
KeePass 1.x CSV Passwords
KeePass 1.x XML Passwords
Large number of US Phone Numbers
Large number of US Zip Codes
Lightweight Directory Access Protocol
Metasploit Module
MySQL database dump
MySqlite database dump
Network Proxy Auto-Config
Nmap Scan Report
PGP Header
PGP Private Key Block
PKCS7 Encrypted Data
Password etc passwd
Password etc shadow
PlainText Private Key
PuTTY SSH DSA Key
PuTTY SSH RSA Key
<table>
<thead>
<tr>
<th>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:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Full names</td>
</tr>
<tr>
<td>• Mailing addresses</td>
</tr>
</tbody>
</table>
• 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
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

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.

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

Topics
- AWS CloudTrail Events (p. 34)
- AWS CloudTrail Errors (p. 34)

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, with 10 being the highest risk and 1 being the lowest.

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.
Viewing Data and Activity that Amazon Macie Monitors

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. 15), 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. 34), 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. 36)
   - S3 objects (p. 36)
   - S3 objects by PII (p. 37)
   - S3 public objects by buckets (p. 38)
   - S3 objects by ACL (p. 38)
   - CloudTrail events and associated users (p. 39)
   - CloudTrail errors and associated users (p. 39)
   - Activity location (p. 40)
   - AWS CloudTrail events (p. 41)
   - Activity ISPs (p. 41)
   - AWS CloudTrail user identity types (p. 41)
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. 54).

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. 54).

## 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. 15).

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. 54).
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. 54).
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. 54).

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.

   ```sql
   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. 54).

**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.

```sql
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. 54).
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. 54).

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. 54).

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. 54).
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 (p. 42)
- Alert Categories in Macie (p. 42)
- Severity Levels for Alerts in Macie (p. 43)
- Locating and Analyzing Macie Alerts (p. 44)
- Adding New and Editing Existing Custom Basic Alerts (p. 45)
- Working with Existing Alerts (p. 46)
- Group Archiving Alerts (p. 46)
- Whitelisting Users or Buckets for Basic Alerts (p. 46)

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. 45).

- **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 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. 43).
- A timestamp that indicates when the alert was generated or last updated.
- The alert category. For more information, see Alert Categories in Macie (p. 42).
- 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. 45).
- 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. 45).
- 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. 54).
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. 42).
     - 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. 54).
     - 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. 43)
     - 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.

   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.

   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,
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. 42).

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 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.
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. 49).

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. 49).
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 (p. 49)
- User Categories in Macie (p. 51)
- Investigating Users (p. 51)

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.

<table>
<thead>
<tr>
<th>userIdentity</th>
<th>MacieUniqueId</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;userIdentity&quot;: {</td>
<td>123456789012:assumed-role/accounting-role</td>
</tr>
<tr>
<td>&quot;type&quot;: &quot;AssumedRole&quot;</td>
<td></td>
</tr>
<tr>
<td>&quot;arn&quot;:</td>
<td></td>
</tr>
<tr>
<td>&quot;arn:aws:sts::123456789012:assumed-role/accounting-role/Mary&quot;</td>
<td></td>
</tr>
<tr>
<td><code>userIdentity</code></td>
<td>MacieUniqueID</td>
</tr>
<tr>
<td>-----------------------------------------------------</td>
<td>--------------------------</td>
</tr>
</tbody>
</table>
| `{                                    
  "type": "IAMUser",
  "arn":
   "arn:aws:iam::123456789012:user/Bob",
  "userName": "Bob"
}`                  | 123456789012:user:bob   |
| `{                                    
  "type": "FederatedUser"
  "arn":
   "arn:aws:sts::123456789012:federated-user/Alice",
  "principalId":
   "123456789012:Alice",
}`                 | 123456789012:federated-user:alice |
| `{                                    
  "type": "AWSAccount"
  "accountId":
   "ANONYMOUS_PRINCIPAL"
}`                  | 123456789012:ANONYMOUS_PRINCIPAL |
| `{                                    
  "type": "Root"
  "sourceARN":
   "arn:aws:iam::123456789012:root",
}`                  | 123456789012:root:root     |
| `{                                    
  "invokedBy":
   "codepipeline.amazonaws.com",
  "type": "AWSService"
}`                 | 123456789012:codepipeline.amazonaws.com |
| `{                                    
  "type": "AWSAccount"
  "accountId":
   "ANONYMOUS_PRINCIPAL"
}`                  | 123456789012:root:root     |
| `{                                    
  "type": "AWSAccount"
  "accountId":
   "987654321098",
  "principalId":
   "AIDABCDEFGH123456XYZ",
}`                 | 123456789012:AIDABCDEFGH123456XYZ |
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. 51)
   - **High-risk CloudTrail errors** (p. 52)
   - **Activity location** (p. 52)
   - **CloudTrail events** (p. 52)
   - **Activity ISPs** (p. 52)
   - **CloudTrail user identity types** (p. 52)

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. 54).

**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. 54).

**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. 54).

**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. 54).

**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. 54).

**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. 54).
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. 54). You can be redirected to the Research tab from various places throughout the Macie console: for example, any of the Dashboard views (see Viewing Data and Activity that Amazon Macie Monitors (p. 35)) or the Basic alerts list (see Amazon Macie Alerts (p. 42)). 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 (p. 54)
- Research Filters (p. 56)
- Saving a Query as an Alert (p. 57)
- Favorite Queries (p. 57)
- Researching AWS CloudTrail Data (p. 57)
- Researching S3 Bucket Properties Data (p. 71)
- Researching S3 Objects Data (p. 79)

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 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

<table>
<thead>
<tr>
<th>Example Query</th>
<th>Description</th>
<th>Data Repository</th>
</tr>
</thead>
<tbody>
<tr>
<td>objectsRead.key:* AND @timestamp:[2017-08-01 TO 2017-12-31]</td>
<td>Search for S3 objects read in the fourth quarter of 2017.</td>
<td>CloudTrail data</td>
</tr>
<tr>
<td>sourceIPAddress.ip_intel.type:&quot;TOR&quot; AND @timestamp:[now-1M TO now]</td>
<td>Search for anonymous accesses to your Macie-monitored data</td>
<td>CloudTrail data</td>
</tr>
<tr>
<td>Example Query</td>
<td>Description</td>
<td>Data Repository</td>
</tr>
<tr>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td><code>macieUniqueId:085924634393\role:malicious_user</code> AND @timestamp: [2018-01-18 TO *]</td>
<td>Search for AWS activities of an assumed role named &quot;malicious_user&quot; in the AWS account ID 085924634393, starting from January 18, 2018.</td>
<td>CloudTrail data</td>
</tr>
</tbody>
</table>

### Example Queries: Integer Field Type

<table>
<thead>
<tr>
<th>Example Query</th>
<th>Description</th>
<th>Data Repository</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>dlp_risk&gt;6</code> AND <code>filesystem_metadata.server_encryption: none</code></td>
<td>Search for S3 objects with a <code>dlp_risk</code> score greater than 6 and without a server-side encryption.</td>
<td>S3 objects</td>
</tr>
<tr>
<td><code>filesystem_metadata.size: [10240 TO 1024000] AND pii_types:*</code></td>
<td>Search for S3 objects between the sizes of 10 MB to 1 GB that contain potential PII data.</td>
<td>S3 objects</td>
</tr>
</tbody>
</table>

### Example Queries: String Field Type

<table>
<thead>
<tr>
<th>Example Query</th>
<th>Description</th>
<th>Data Repository</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>dlp_risk&gt;5</code> AND `key: /.<em>contract.</em></td>
<td>.<em>agreement.</em></td>
<td>.*terms./` AND @timestamp:[now-1M/M TO now]</td>
</tr>
</tbody>
</table>

**Note**
Some regex queries might result in long search times. We recommend conducting searches for limited time frames.

<table>
<thead>
<tr>
<th>Example Query</th>
<th>Description</th>
<th>Data Repository</th>
</tr>
</thead>
<tbody>
<tr>
<td>`mimetypes:&quot;Adobe PDF (application/pdf)&quot; AND key: /-(.*.pdf</td>
<td>.*.PDF)/`</td>
<td>Search for S3 objects containing PDF data but in files with file extensions other than PDF/pdf.</td>
</tr>
</tbody>
</table>

**Note**
This query also returns archived objects (zip,7z, etc.) containing PDF documents.
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

---

<table>
<thead>
<tr>
<th>Example Query</th>
<th>Description</th>
<th>Data Repository</th>
</tr>
</thead>
<tbody>
<tr>
<td>acl.Grants.Grantee.DisplayName: admin</td>
<td>Search for S3 buckets with ACL grantee display names set to &quot;admin.&quot;</td>
<td>S3 bucket properties</td>
</tr>
<tr>
<td>acl.Grants.Grantee.DisplayName: admi?</td>
<td>Search for S3 buckets with ACL grantee display names set to &quot;admi(?)&quot; (wildcard), including &quot;admin.&quot;</td>
<td>S3 bucket properties</td>
</tr>
<tr>
<td>bucket: <em>test</em></td>
<td>Search for S3 buckets with keywords &quot;test.&quot;</td>
<td>S3 bucket properties</td>
</tr>
</tbody>
</table>
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. 42).

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. 45).

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 (p. 57)
- CloudTrail Data Fields and Sample Queries (p. 58)

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.

<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>CloudTrail Field Name</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
<tbody>
<tr>
<td>accountId</td>
<td>userIdentity.accountId</td>
<td>String</td>
<td>The AWS account ID.</td>
<td>search for user sessions with accesses related to a particular account: accountid:“110912345678”</td>
</tr>
<tr>
<td>Macie Field Name</td>
<td>CloudTrail Field Name</td>
<td>Macie Field Type</td>
<td>Description</td>
<td>Example Search Query</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------</td>
<td>------------------</td>
<td>-------------</td>
<td>----------------------</td>
</tr>
</tbody>
</table>
| 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           | String           | The event name that resulted in the returned (if any) error code. | • Search for user sessions with any AWS ConsoleLogin call:  
  • eventNameErrorCode.key:ConsoleLogin  
  • Search for user sessions with any AWS Delete call:  
  • eventNameErrorCode.key:Delete* |
| eventNameErrorCode.secondary | secondaryErrorCode | 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:"AccessDenied" |
| 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.com"  
  • eventSource.key:"lambda.amazonaws.com" |
| 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" |
<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>CloudTrail Field Name</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
</table>
| 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. | 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:"DeleteObject:AccessDenied"`  
  
  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:
<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>CloudTrail Field Name</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>objectsDeleted.key:/</td>
</tr>
<tr>
<td>Macie Field Name</td>
<td>CloudTrail Field Name</td>
<td>Macie Field Type</td>
<td>Description</td>
<td>Example Search Query</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------</td>
<td>------------------</td>
<td>-------------</td>
<td>----------------------</td>
</tr>
</tbody>
</table>
| 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:"GetObject:AccessDenied"  
  
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:  


<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>CloudTrail Field Name</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
<tbody>
<tr>
<td>objectsRead.key:/arn:aws:s3:::my_sensitive_bucket.*</td>
<td>objectsRead.key</td>
<td>Key</td>
<td></td>
<td>objectsRead.key:/arn:aws:s3:::my_sensitive_bucket.* / AND accountid:&quot;ANONYMOUS_PRINCIPAL&quot;</td>
</tr>
<tr>
<td>Macie Field Name</td>
<td>CloudTrail Field Name</td>
<td>Macie Field Type</td>
<td>Description</td>
<td>Example Search Query</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------</td>
<td>-----------------</td>
<td>-------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>objectsWritten.key</td>
<td>Resources[0].ARN</td>
<td>String</td>
<td>A list of S3 object ARNs that were part of a PutObject, CopyObject, or CompleteMultipartUpload API call.</td>
<td>Search for user sessions with all objects written to a particular bucket: <code>objectsWritten.key:/arn:aws:s3:::my_bucket_name/*</code></td>
</tr>
</tbody>
</table>

**Note**
- Objects that are part of a failed PutObject API call are also added to the aggregate record of objectsWritten.key.
- A user session that returns the results of a search against objectsWritten.key has a maximum limit of 250 records.

Additional search queries:
- 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 event and any attempt to read a particular sensitive object, bucket, or prefix: `objectsWritten.key:/arn:aws:s3:::my_sensitive_bucket/* AND eventNameErrorCode.compound:"PutObject:AccessDenied"`
- 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:
<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>CloudTrail Field Name</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
</table>
| principalId      | principalId            | String          | The IAM principal ID. | Search for user sessions with access requests from a particular principal ID:  
|                  |                        |                 | Note When an assumed role makes a request, the session name is removed from the principal ID. | principalId:"AIDAIMABCKFJSKEOAKWNE" |
| recipientAccountId| recipientAccountId     | String          | The account ID that received the CloudTrail event. | Search for all activity in a particular account:  
|                  |                        |                 | | recipientAccountId:"110912345678" |
| resourceOwnerAccountIds.key | resourceOwnerAccountIds.key | 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].accountId | 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_PRINCIPAL" |

- objectsWritten.key:/arn:aws:s3:::my_sensitive_bucket/*,
- AND accountid:"ANONYMOUS_PRINCIPAL"
<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>CloudTrail Field Name</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
</table>
| 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.String    | 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::123456789123:cluster-api" |
<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>CloudTrail Field Name</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
</table>
| sourceIPAddress.key | sourceIPAddress | String | The IP address that the request was made from. | 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.RunInstances_count: >10  AND NOT (sourceIPAddress.key: "autoscaling.amazonaws.com") |
| 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 | 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.
<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
<tbody>
<tr>
<td>countLongLifeAccessToken</td>
<td>Integer</td>
<td>A count of GetSessionToken API calls with a lifespan longer than the default 43,200 seconds.</td>
<td>• @timestamp:=&quot;2017-02-06&quot; &lt;br&gt;Search for user sessions with access requests between two time intervals: &lt;br&gt;• @timestamp:[2017-02-01 TO 2017-02-27]</td>
</tr>
<tr>
<td>dcObjectsDeleted</td>
<td>Integer</td>
<td>A count of unique S3 objects deleted in a user session. &lt;br&gt;Note &lt;br&gt;A user session that returns the results of a search against dcObjectsDeleted has a maximum limit of 250 entries.</td>
<td>Search for user sessions with more than 25 distinct objects deleted by an AWS user or a role: &lt;br&gt;• dcObjectsDeleted:&gt;25 &lt;br&gt;• dcObjectsDeleted:[25 TO 100]</td>
</tr>
<tr>
<td>dcObjectsRead</td>
<td>Integer</td>
<td>A count of unique S3 objects read in a user session. &lt;br&gt;Note &lt;br&gt;A user session that returns the results of a search against dcObjectsRead has a maximum limit of 250 entries.</td>
<td>Search for user sessions with more than 25 distinct objects read by an AWS user or a role: &lt;br&gt;• dcObjectsRead:&gt;25 &lt;br&gt;• dcObjectsRead:[25 TO 100] &lt;br&gt;Search for more than 25 distinct objects read by an anonymous principal during a user session: &lt;br&gt;• dcObjectsRead:&gt;25 AND accountid:&quot;ANONYMOUS_PRINCIPAL&quot;</td>
</tr>
<tr>
<td>dcObjectsWritten</td>
<td>Integer</td>
<td>A count of unique S3 objects written in a user session. &lt;br&gt;Note &lt;br&gt;A user session that returns the results of a search against dcObjectsWritten has a maximum limit of 250 entries.</td>
<td>Search for user sessions with more than 25 distinct objects written by an AWS user or a role: &lt;br&gt;• dcObjectsWritten:&gt;25 &lt;br&gt;• dcObjectsWritten:[25 TO 100]</td>
</tr>
<tr>
<td>Macie Field Name</td>
<td>Macie Field Type</td>
<td>Description</td>
<td>Example Search Query</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>distinctEventName</td>
<td>Integer</td>
<td>A count of unique event names that take place in a user session.</td>
<td>Search for user sessions with more than 25 unique API calls being executed by a user or a role:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- distinctEventName: &gt;25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- distinctEventName: [25 TO 100]</td>
</tr>
<tr>
<td>distinctSourceIPAddress</td>
<td>Integer</td>
<td>A count of unique source IP addresses involved in activity that takes place in a user session. The maximum value is 60,000.</td>
<td>Search for user sessions with more than 25 distinct source IP addresses observed for a user or a role:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- distinctSourceIPAddress: &gt;25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- distinctSourceIPAddress: [25 TO 100]</td>
</tr>
<tr>
<td>distinctUserAgent</td>
<td>Integer</td>
<td>A count of unique client user agents involved in activity that takes place in a user session. The maximum value is 60,000.</td>
<td>Search for user sessions with more than 25 user agents observed for a user or a role:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- distinctUserAgent: &gt;25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- distinctUserAgent: [25 TO 100]</td>
</tr>
<tr>
<td>eventNameErrorCode.compound</td>
<td>String compound</td>
<td>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.</td>
<td>Search for user sessions with AccessDenied error while attempting a GetObject call:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- eventNameErrorCode.compound: &quot;GetObject:AccessDenied&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Search for user sessions with any errors associated with PutObject calls:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- eventNameErrorCode.compound: / PutObject:.* /</td>
</tr>
<tr>
<td>eventNameIsp.com</td>
<td>String</td>
<td>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.</td>
<td>Search for user sessions with ConsoleLogin calls from non-AWS IPs using a regular expression:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- eventNameIsp.com.compound: / ConsoleLogin: ~(Amazon.*) /</td>
</tr>
<tr>
<td>Macie Field Name</td>
<td>Macie Field Type</td>
<td>Description</td>
<td>Example Search Query</td>
</tr>
<tr>
<td>------------------------</td>
<td>------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>eventNameIsp.secondary</td>
<td>String</td>
<td>The ISP that the AWS API call was made from.</td>
<td>Search for user sessions with AWS API calls coming from outside Amazon IP addresses:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• eventNameIsp.secondary:/~(Amazon.*)/</td>
</tr>
<tr>
<td>macieUniqueId</td>
<td>String</td>
<td>A format that is unique to Macie for specifying users. The Macie unique ID is a combination of the IAM Identity and the recipientAccountId. For more information, see MacieUniqueID (p. 49).</td>
<td>Search for user sessions with accesses from a particular role, user, or root account:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• macieUniqueId:&quot;123456789123:assumed-role:co-admin&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• macieUniqueId:&quot;123456789123:root:root&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• macieUniqueId:&quot;123456789123:user:exampleuser&quot;</td>
</tr>
<tr>
<td>sourceIPAddress.ip_intel.type</td>
<td>String</td>
<td>The IP intelligence category associated with a source IP address.</td>
<td>Search for user sessions with all accesses from a Tor network:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• sourceIPAddress.ip_intel.type:&quot;TOR&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Search for user sessions with all accesses from threat intelligence input feeds:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• sourceIPAddress.ip_intel.type:*</td>
</tr>
<tr>
<td>windowStartTimeInMillis</td>
<td>Integer</td>
<td>The epoch timestamp for the start of a user session.</td>
<td>Search for user sessions whose first event time is greater than a given epoch time:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• windowStartTimeInMillis:&gt;1424476529</td>
</tr>
<tr>
<td>windowEndTimeInMillis</td>
<td>Integer</td>
<td>The epoch timestamp for the end of a user session.</td>
<td>Search for user sessions whose last event time is less than a given epoch time:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• windowEndTimeInMillis:&lt;1424476987</td>
</tr>
<tr>
<td>ipLocation.key</td>
<td>String</td>
<td>The IP geolocation (city and country) accessed by an identity that Macie monitors.</td>
<td>Search for user sessions with any AWS API call events originating in Los Angeles:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• ipLocation.key:&quot;LosAngeles:UnitedStates&quot;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Search for user session any AWS API call events originating from outside the United States:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• ipLocation.key:/~(.*UnitedStates)/</td>
</tr>
</tbody>
</table>
### Researching S3 Bucket Properties Data

#### Topics
- Analyzing S3 Buckets Properties Search Results (p. 71)
- S3 Bucket Properties Data Fields and Example Queries (p. 72)

#### 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.

<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
</table>
| 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.*)/` |
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

<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>Amazon S3 API Field Name</th>
<th>Amazon S3 API Operation</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
</table>
| acl.Grants.Grantee.DisplayName | Grants.Grantee.DisplayName | getbucket-acl | String | The display name of the S3 bucket ACL grantee. | Search for S3 buckets accessible by John Doe:  
- acl.Grants.Grantee.DisplayName:="JohnDoe" |
| acl.Grants.Grantee.ID | Grants.Grantee.ID | getbucket-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:="75bee88dfe7bf9b7243aea123934be3e55d777d8412b4aa1703284e2bb29371b" |
| acl.Grants.Grantee.Type | Grants.Grantee.Type | getbucket-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:CanonicalUser |
<p>| acl.Grants.Grantee.URI | Grants.Grantee.URI | getbucket-acl | String | The URI identifier of the S3 bucket ACL grantee. | Search for all S3 buckets except those that |</p>
<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>Amazon S3 API Field Name</th>
<th>Amazon S3 API Operation</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
<tbody>
<tr>
<td>acl.Grants.Grantee.URI</td>
<td>acl.Grants.Grantee.URI</td>
<td>bucket-acl</td>
<td>String</td>
<td>The permission level assigned to the ACL grantee.</td>
<td>Search for S3 buckets that grant full (read/write) access to anyone:</td>
</tr>
<tr>
<td>acl.Grants.Grantee.URI</td>
<td>acl.Grants.Grantee.URI</td>
<td>bucket-acl</td>
<td>String</td>
<td>The display name of the S3 bucket owner.</td>
<td>Search for S3 buckets owned by John Doe:</td>
</tr>
<tr>
<td>acl.Owner.DisplayName</td>
<td>acl.Owner.DisplayName</td>
<td>bucket-acl</td>
<td>String</td>
<td>The ID of the S3 bucket owner.</td>
<td>Search for a particular S3 bucket owner ID:</td>
</tr>
</tbody>
</table>

For example:

- Search for S3 buckets that belong to the LogDelivery group:
  

- 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"

- Search for S3 buckets that grant full (read/write) access to anyone:
  

- Search for S3 buckets owned by John Doe:
  
  • acl.Owner.DisplayName:"JohnDoe"

- Search for a particular S3 bucket owner ID:
  
  • acl.Owner.ID:"73bee78dfe7b89b7243aea1c6934baae55d777d8412b4fd1703284e2bb29371f"
<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>Amazon S3 API Field Name</th>
<th>Amazon S3 API Operation</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
<tbody>
<tr>
<td>location.LocationConstraint</td>
<td>LocationConstraint</td>
<td>get-bucket-location</td>
<td>String</td>
<td>The AWS Region where the S3 bucket resides. <strong>Note</strong> 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”.</td>
<td>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”</td>
</tr>
<tr>
<td>logging.LoggingEnabled.TargetBucket</td>
<td>LoggingEnabled.TargetBucket</td>
<td>get-bucket-logging</td>
<td>String</td>
<td>The bucket whose logging status is being returned.</td>
<td>Search for all buckets with S3 object level logging enabled: • logging.LoggingEnabled.TargetBucket:*</td>
</tr>
<tr>
<td>logging.LoggingEnabled.TargetPrefix</td>
<td>LoggingEnabled.TargetPrefix</td>
<td>get-bucket-logging</td>
<td>String</td>
<td>The configured prefix or folder containing Object Level Logging data for a particular S3 bucket.</td>
<td>Search for buckets configured with a prefix substring of “Production”: • logging.LoggingEnabled.TargetPrefix:”Production”</td>
</tr>
<tr>
<td>policy.Policy.Id</td>
<td>Policy.Id</td>
<td>get-bucket-policy</td>
<td>String</td>
<td>The ID for an S3 bucket policy.</td>
<td>Search for bucket policies with a particular ID: • policy.Policy.Id:”aaaa-bbbb-cccc-dddd”</td>
</tr>
<tr>
<td>Macie Field Name</td>
<td>Amazon S3 API Field Name</td>
<td>Amazon S3 API Operation</td>
<td>Macie Field Type</td>
<td>Description</td>
<td>Example Search Query</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------------</td>
<td>-------------------------</td>
<td>------------------</td>
<td>-------------</td>
<td>----------------------</td>
</tr>
</tbody>
</table>
| policy.Policy.Statement.Effect          | Policy.Statement.Effect | get-bucket-policy       | String           | The list of policy effects associated with an S3 bucket policy. | Search for bucket policies with explicit "allow" grants:  
| policy.Policy.Statement.NotPrincipal.AWS | Policy.Statement.NotPrincipal.AWS | get-bucket-policy | 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.NotPrincipal.CanonicalUser: 79a59df900b949e55d96a1e698fbacedfd6e09d98eacf8f8d5218e7cd47ef2be` |
<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>Amazon S3 API Field Name</th>
<th>Amazon S3 API Operation</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macie Field Name</td>
<td>Amazon S3 API Field Name</td>
<td>Amazon S3 API Operation</td>
<td>Macie Field Type</td>
<td>Description</td>
<td>Example Search Query</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------------</td>
<td>-------------------------</td>
<td>------------------</td>
<td>-------------</td>
<td>----------------------</td>
</tr>
</tbody>
</table>
| 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.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:"2012-10-17" |
| 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" |
<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>Amazon S3 API Field Name</th>
<th>Amazon S3 API Operation</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
</table>
| website.IndexDocument.Suffix | IndexDocument.Suffix | get-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" |

For more information, see GET Bucket lifecycle.

• 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.NoncurrentVersionTransitions.StorageClass
• lifecycle_configuration.Rules.Prefix
• lifecycle_configuration.Rules.Transitions.Date
• lifecycle_configuration.Rules.Transitions.Days
• lifecycle_configuration.Rules.Transitions.StorageClass

Search for the S3 buckets with a lifecycle configuration rule whose expiration is less than 3 days:

• lifecycle_configuration.Rules.Expiration.Days:<3

S3 Bucket Properties Data Fields That Macie Generates

<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example search query</th>
</tr>
</thead>
</table>
| @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: |
Researching S3 Objects Data

<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example search query</th>
</tr>
</thead>
</table>
| 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" | 79

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

<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>Amazon S3 API Field Name</th>
<th>Amazon S3 API Operation</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
<tbody>
<tr>
<td>key</td>
<td>key</td>
<td>get-bucket (listObjects)</td>
<td>String</td>
<td>The S3 object key path.</td>
<td><strong>Search for document names with the keyword “myobject”:</strong>&lt;br&gt;• <code>key: /.*/myobject.*</code></td>
</tr>
<tr>
<td>accountId</td>
<td>None</td>
<td>None</td>
<td>String</td>
<td>The AWS account ID that owns the S3 object.</td>
<td><strong>Search for S3 objects owned by a particular account ID:</strong>&lt;br&gt;• <code>accountId:&quot;110912345678&quot;</code></td>
</tr>
<tr>
<td>filesystem_metadata.bucket</td>
<td>None</td>
<td>None</td>
<td>String</td>
<td>The S3 bucket name that holds the S3 object.</td>
<td><strong>Search for S3 objects in a particular S3 bucket:</strong>&lt;br&gt;• <code>filesystem_metadata.bucket:&quot;MyBucket&quot;</code></td>
</tr>
<tr>
<td>Macie Field Name</td>
<td>Amazon S3 API Field Name</td>
<td>Amazon S3 API Operation</td>
<td>Description</td>
<td>Example Search Query</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------</td>
<td>-------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>filesystem_metadata.first_prefix</td>
<td>get- prefix bucket (listObjects)</td>
<td>String</td>
<td>The name of the first folder that contains the S3 object.</td>
<td>Search for S3 objects contained in first folder names where folder name is AWSLogs:</td>
<td></td>
</tr>
<tr>
<td>Note</td>
<td></td>
<td></td>
<td>Macie uses the S3 key field and parses out everything before the first '/,' not including the bucket name.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>filesystem_metadata.ETag</td>
<td>get-bucket (listBuckets)</td>
<td>String</td>
<td>The entity tag as defined in RFC 2616.</td>
<td>Search for a particular eTag:</td>
<td></td>
</tr>
<tr>
<td>filesystem_metadata.bucket_owner.id</td>
<td>get-bucket-acl</td>
<td>String</td>
<td>The unique ID of the S3 bucket owner.</td>
<td>Search for S3 objects belonging to a particular owner ID:</td>
<td></td>
</tr>
<tr>
<td>filesystem_metadata.bucket_owner.name</td>
<td>get-bucket-acl</td>
<td>String</td>
<td>The name of the S3 bucket owner.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>filesystem_metadata.last_modified</td>
<td>get-bucket (list-buckets)</td>
<td>Date</td>
<td>The timestamp when the S3 object was last modified.</td>
<td>Search for S3 objects that were modified in the last 24 hours:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• filesystem_metadata.last_modified: [now-1d TO now]</td>
<td></td>
</tr>
<tr>
<td>Macie Field Name</td>
<td>Amazon S3 API Field Name</td>
<td>Amazon S3 API Operation</td>
<td>Macie Field Type</td>
<td>Description</td>
<td>Example Search Query</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------------</td>
<td>------------------------</td>
<td>------------------</td>
<td>-------------</td>
<td>---------------------</td>
</tr>
</tbody>
</table>
| filesystem_metadata.server_encryption | 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_encryption: "AES256" |
| filesystem_metadata.size | Size | 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 | 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_key_id: "arn:aws:kms:us-west-2:110912345678:key/06f8b4fa-3f50-4aad-8a98-b60a56a9a1f2" |
| object_acl.Grants.Grantee.DisplayName | 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.DisplayName: "JohnDoe" |
| object_acl.Grants.Grantee.ID | 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: "75aa57f09aa0c8caeab4f8c24e99d10f8e7faeebf76c078efc7c6caea54ba06a" |
| object_acl.Grants.Grantee.Type | 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: "CanonicalUser"  
• object_acl.Grants.Grantee.Type: "Group" |
## S3 Objects Data Fields and Sample Queries

<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>Amazon S3 API Field Name</th>
<th>S3 API Operation</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
</table>
| object_acl.Grants.Grantee.URI | object_acl.Grants.Grantee.URI | GET object-acl | String | The ACL grantee URI. | Search for S3 object ACLs with the AllUsers grant: 
| object_acl.Grants.Permission | object_acl.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 | object_acl.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 | object_acl.Owner.ID | GET object-acl | String | The ACL owner ID. | Search for S3 objects belonging to a particular owner ID: 
  - object_acl.Owner.ID: "447fba12b05da301df359096ff54dd86bbbe347d3d5aff1e59f0906cd6a8394d2" |

### S3 Object Data Fields That Macie Generates

<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
</table>
| @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" |
<table>
<thead>
<tr>
<th>Macie Field Name</th>
<th>Macie Field Type</th>
<th>Description</th>
<th>Example Search Query</th>
</tr>
</thead>
<tbody>
<tr>
<td>console's Settings page.</td>
<td></td>
<td></td>
<td>content_type:&quot;text/x-java-source&quot; AND regex_themes: &quot;aws_access_key&quot;</td>
</tr>
<tr>
<td>dlp_risk</td>
<td>Integer</td>
<td>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.</td>
<td>Search for globally accessible (read or write) objects with the compound (final) risk level of 5 or higher: object_acl.Grants.Grantee.URI: &quot;<a href="http://acs.amazonaws.com/groups/global/AllUsers">http://acs.amazonaws.com/groups/global/AllUsers</a>&quot; AND dlp_risk&gt;5</td>
</tr>
<tr>
<td>encoding</td>
<td>String</td>
<td>The encoding scheme identified when analyzing the S3 object content.</td>
<td>Search for Unicode text documents: encoding: &quot;utf-8&quot;</td>
</tr>
<tr>
<td>filetype_risk</td>
<td>Integer</td>
<td>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.</td>
<td>Search for documents with the assigned file extension risk of greater than 6: filetype_risk: &gt; 6</td>
</tr>
<tr>
<td>Macie Field Name</td>
<td>Macie Field Type</td>
<td>Description</td>
<td>Example Search Query</td>
</tr>
<tr>
<td>---------------------</td>
<td>------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>filetypes</td>
<td>String</td>
<td>The type of the file based on the extension.</td>
<td>Search for files with an extension of .pdf:</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Note</strong> 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.</td>
<td>• filetypes: &quot;Adobe PDF (.pdf)&quot;</td>
</tr>
<tr>
<td>keyword_themes</td>
<td>String</td>
<td>The themes assigned to an S3 object. You can find supported themes in the Macie console's Settings page.</td>
<td>Search for S3 objects containing content related to Social Security:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• keyword_themes: &quot;Social Security Keywords&quot;</td>
</tr>
<tr>
<td>language_code</td>
<td>String</td>
<td>The language code found when analyzing the S3 object's content.</td>
<td>Search for S3 objects containing German keywords:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• language_code: &quot;de&quot;</td>
</tr>
<tr>
<td>last_crawl_time</td>
<td>Date</td>
<td>The timestamp of when Macie last analyzed an S3 object.</td>
<td>Search for S3 objects analyzed by Macie in the last 24 hours:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• last_crawl_time: [now-1d/d TO now]</td>
</tr>
<tr>
<td>mimetype_risk</td>
<td>Integer</td>
<td>The risk level based on an S3 object's content / MIME type.</td>
<td>Search for S3 objects containing MIME types associated with high-risk content:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• mimetype_risk: &gt; 5</td>
</tr>
<tr>
<td>mimetypes</td>
<td>String</td>
<td>The MIME type of an S3 object.</td>
<td>Search for plaintext documents containing AWS secret keys:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• mimetypes: &quot;Plain Text (text/plain)&quot; AND themes: aws_secret_key</td>
</tr>
<tr>
<td>pii_impact</td>
<td>String</td>
<td>The PII severity impact of an S3 object, assigned by Macie.</td>
<td>Search for S3 objects containing highly valuable personal identifiable information:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• pii_impact: &quot;high&quot;</td>
</tr>
<tr>
<td>Macie Field Name</td>
<td>Macie Field Type</td>
<td>Description</td>
<td>Example Search Query</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------</td>
<td>-------------</td>
<td>----------------------</td>
</tr>
</tbody>
</table>
| 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" |
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 re-enable 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. 42).

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).

Event Format

The event for Macie in CloudWatch Events has the following format. The fictional account ID "111122223333" represents the ID of the master Macie account.

```json
{
    "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": [
    ],
    "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",
        "risk-score": 8,
        "trigger": {
            "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": "55556666777:assumed-role:superawesome:aroaidpldc7nesfnheji",
        "summary": {ALERT_DETAILS_JSON
    }
}
```
Configure CloudWatch Events

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.

**Prerequisite**

Create an Amazon SQS queue for the events from Macie. For more information, see Tutorial: Creating an Amazon SQS Queue.

**To configure CloudWatch events for your master Macie account**

2. In the navigation pane, choose Events, Rules and then choose Create rule.
3. Choose Edit and enter the following event pattern for the Macie events.

   ```json
   { "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.
Document History for Amazon Macie

The following table describes the updates for Amazon Macie.

<table>
<thead>
<tr>
<th>update-history-change</th>
<th>update-history-description</th>
<th>update-history-date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adds support for service-linked roles for Macie</td>
<td>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.</td>
<td>June 28, 2018</td>
</tr>
<tr>
<td>Adds descriptions of data fields that can appear in the results of your data searches (p. 90)</td>
<td>For more information, see CloudTrail Data Fields, S3 Bucket Properties Data Fields, and S3 Objects Data Fields.</td>
<td>May 4, 2018</td>
</tr>
<tr>
<td>Initial release (p. 90)</td>
<td>Initial release of the <em>Amazon Macie User Guide</em></td>
<td>August 14, 2017</td>
</tr>
</tbody>
</table>