

---

# Amazon Macie

## User Guide



## **Amazon Macie: User Guide**

Copyright © 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
Accessing Amazon Macie	3
Pricing for Amazon Macie	3
Related services	4
Getting started	5
Before you begin	5
Step 1: Enable Amazon Macie	5
Step 2: Configure a repository for sensitive data discovery results	6
Step 3: Create a job to discover sensitive data	6
Step 4: Review your findings	7
Monitoring Amazon S3 data	8
How Macie monitors Amazon S3 data	8
Key components	9
Data refreshes	10
Additional considerations	11
Assessing your Amazon S3 security posture	12
Displaying the dashboard	13
Understanding dashboard components	13
Understanding S3 bucket statistics on the dashboard	15
Analyzing your Amazon S3 security posture	17
Reviewing your S3 bucket inventory	18
Filtering your S3 bucket inventory	24
Allowing Macie to access S3 buckets and objects	32
Discovering sensitive data	36
Using managed data identifiers	37
Keyword requirements	37
Sensitive data categories and types	38
Building custom data identifiers	53
Components of a custom data identifier	54
Creating custom data identifiers	55
Regex support in custom data identifiers	56
Running sensitive data discovery jobs	56
Scope options for jobs	57
Creating a job	65
Monitoring jobs	71
Reviewing job statistics and results	80
Managing jobs	82
Forecasting and monitoring job costs	87
Supported file and storage formats	89
Analyzing encrypted S3 objects	90
Encryption options for S3 objects	91
Allowing Macie to use a customer managed KMS key	92
Storing and retaining sensitive data discovery results	96
Step 1: Verify your permissions	96
Step 2: Define the AWS KMS key and policy	97
Step 3: Specify the S3 bucket to use	99
Troubleshooting errors	103
Analyzing findings	104
Types of findings	105
Policy findings	105
Sensitive data findings	106
Viewing findings	107
Locating sensitive data with findings	108

---

Locating sensitive data .....	109
Schema for sensitive data locations .....	110
Schema details and examples .....	112
Filtering findings .....	117
Filter fundamentals .....	118
Creating and applying filters .....	123
Creating and managing filter rules .....	129
Fields for filtering findings .....	134
Suppressing findings .....	152
Creating suppression rules .....	152
Viewing suppressed findings .....	154
Changing suppression rules .....	155
Deleting suppression rules .....	156
Severity scoring for findings .....	157
Severity scoring for policy findings .....	158
Severity scoring for sensitive data findings .....	158
Monitoring and processing findings .....	162
EventBridge integration .....	162
Using EventBridge .....	163
Creating EventBridge rules for finding events .....	163
Security Hub integration .....	166
How Macie publishes findings to Security Hub .....	166
Examples of Macie findings in Security Hub .....	169
Enabling and configuring Security Hub integration .....	173
Stopping the publication of findings to Security Hub .....	173
Configuring publication settings for findings .....	173
Choosing publication destinations .....	174
Determining the publication frequency .....	174
Changing the publication frequency .....	175
EventBridge event schema for findings .....	175
Event schema .....	176
Event example for a policy finding .....	176
Event example for a sensitive data finding .....	179
Managing multiple accounts .....	184
Managing multiple accounts through AWS Organizations .....	184
Managing multiple accounts by invitation .....	184
Understanding the relationship between administrator and member accounts .....	184
Managing accounts with AWS Organizations .....	186
Designating a delegated Macie administrator for an AWS organization .....	186
Adding existing organization accounts as members .....	188
Managing accounts by invitation .....	188
Adding a member account .....	189
Inviting an account .....	190
Accepting an invitation .....	190
Logging API calls .....	191
Macie information in CloudTrail .....	191
Understanding Macie log file entries .....	192
Forecasting and monitoring costs .....	193
Understanding how estimated usage costs are calculated .....	193
Reviewing estimated usage costs .....	194
Reviewing estimated usage costs on the Amazon Macie console .....	195
Querying estimated usage costs programmatically with the Amazon Macie API .....	195
Participating in the free trial .....	198
Suspending or disabling Macie .....	200
Suspending Macie .....	200
Disabling Macie .....	201
Security .....	202

---

Data protection .....	202
Encryption at rest .....	203
Encryption in transit .....	203
Identity and access management .....	203
Policy structure .....	203
AWS managed policies .....	204
API actions .....	204
Service-linked roles .....	205
Service-linked role permissions for Macie .....	205
Create a service-linked role for Macie .....	207
Edit a service-linked role for Macie .....	207
Delete a service-linked role for Macie .....	207
AWS managed policies .....	207
Policy updates .....	208
Logging and monitoring .....	208
Compliance validation .....	208
Resilience .....	209
Infrastructure security .....	209
Quotas .....	210
AWS glossary .....	212
Document history .....	213

# What is Amazon Macie?

Amazon Macie is a fully managed data security and data privacy service that uses machine learning and pattern matching to help you discover, monitor, and protect sensitive data in your AWS environment.

Macie automates the discovery of sensitive data, such as personally identifiable information (PII) and financial data, to provide you with a better understanding of the data that your organization stores in Amazon Simple Storage Service (Amazon S3). Macie also provides you with an inventory of your S3 buckets, and it automatically evaluates and monitors those buckets for security and access control. Within minutes, Macie can identify and report overly permissive or unencrypted buckets for your organization.

If Macie detects sensitive data or potential issues with the security or privacy of your data, it creates detailed findings for you to review and remediate as necessary. You can review and analyze these findings directly in Macie, or monitor and process them by using other services, applications, and systems.

## Topics

- [Features of Amazon Macie \(p. 1\)](#)
- [Accessing Amazon Macie \(p. 3\)](#)
- [Pricing for Amazon Macie \(p. 3\)](#)
- [Related services \(p. 4\)](#)

## Features of Amazon Macie

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

### Automate the discovery of sensitive data

With Macie, you can automate discovery and reporting of sensitive data by [creating and running sensitive data discovery jobs \(p. 56\)](#). A sensitive data discovery job analyzes objects in S3 buckets to determine whether they contain sensitive data. If Macie detects sensitive data in an object, it creates a sensitive data finding for you. The finding provides a detailed report of the sensitive data that Macie found.

You can configure a job to run only once, for on-demand analysis and assessment, or on a recurring basis for periodic analysis, assessment, and monitoring. You can also choose various options to control the breadth and depth of a job's analysis—the S3 buckets to analyze, the sampling depth, and custom include and exclude criteria that derive from properties of S3 objects. With these scheduling and scope options, you can build and maintain a comprehensive view of the data that your organization stores in Amazon S3 and any security or compliance risks for that data.

### Discover a variety of sensitive data types

When you run a sensitive data discovery job, the job can use built-in criteria and techniques, such as machine learning and pattern matching, to analyze objects in S3 buckets. These criteria and techniques, referred to as [managed data identifiers \(p. 37\)](#), can detect a large and growing list of sensitive data types for many countries and regions, including multiple types of personally identifiable information (PII), personal health information (PHI), and financial data.

The job can also use [custom data identifiers \(p. 53\)](#) that you create. A custom data identifier is a set of criteria that you define—a regular expression (*regex*) that defines a text pattern to match

and, optionally, character sequences and a proximity rule that refine the results. With this type of identifier, you can detect sensitive data that reflects your particular scenarios, intellectual property, or proprietary data, and supplement the managed data identifiers that Macie provides.

### Evaluate and monitor data for security and access control

When you enable Macie, Macie immediately generates and begins maintaining a complete inventory of your S3 buckets, and it begins evaluating and monitoring the buckets for security and access control. If Macie detects a potential issue with the security or privacy of the data, it creates a [policy finding \(p. 105\)](#) for you.

In addition to specific findings, a [dashboard \(p. 12\)](#) gives you a snapshot of aggregated statistics for your buckets. This includes statistics that indicate how many of your buckets are publicly accessible, are shared with other AWS accounts, or don't encrypt objects by default. You can drill down on each statistic to view the supporting data.

Macie also provides you with detailed information and statistics for individual buckets in your inventory. This data includes breakdowns of a bucket's public access and encryption settings, and the size and number of objects that Macie can analyze to detect sensitive data in the bucket. You can [browse the inventory \(p. 17\)](#), or sort and filter the inventory by certain fields. When you choose a bucket, a panel displays the bucket's details.

### Review and analyze findings

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

To [review, analyze, and manage findings \(p. 104\)](#), you can use the **Findings** pages on the Amazon Macie console. These pages list your findings and provide the details of individual findings. They also provide multiple options for grouping, filtering, sorting, and suppressing findings. You can also use the Amazon Macie API to query, retrieve, and suppress findings. If you use the API, you can pass the data to another application, service, or system for deeper analysis, long-term storage, or reporting.

### Monitor and process findings with other services and systems

To support integration with other services and systems, Macie [publishes findings to Amazon EventBridge \(p. 162\)](#) as finding events. EventBridge is a serverless event bus service that can route findings data to targets such as AWS Lambda functions and Amazon Simple Notification Service (Amazon SNS) topics. With EventBridge, you can monitor and process findings in near-real time as part of your existing security and compliance workflows.

You can configure Macie to also [publish findings to AWS Security Hub \(p. 166\)](#). Security Hub is a service that provides a comprehensive view of your security posture across your AWS environment and helps you check your environment against security industry standards and best practices. With Security Hub, you can more easily monitor and process your findings as part of a broader analysis of your organization's security posture in AWS.

### Centrally manage multiple Macie accounts

If your AWS environment has multiple accounts, you can [centrally manage multiple Macie accounts \(p. 184\)](#) as a single organization. You can do this by using AWS Organizations or by sending membership invitations from Macie.

In a multiple-account configuration, a single Macie administrator can perform certain tasks and manage certain settings for accounts that are members of the same organization. Tasks include viewing information about S3 buckets that are owned by member accounts, viewing policy findings for those buckets, and running sensitive data discovery jobs to detect sensitive data in those buckets. If the accounts are associated through AWS Organizations, the Macie administrator can also enable Macie for member accounts in the organization.

### Develop and manage resources programmatically

In addition to the Amazon Macie console, you can interact with Macie by using the [Amazon Macie API](#). The Amazon Macie API gives you comprehensive, programmatic access to your Macie account and resources.

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

## Accessing Amazon Macie

Amazon Macie is available in most AWS Regions. For a list of Regions where Macie is currently available, see [Amazon Macie endpoints and quotas](#) in the *Amazon Web Services General Reference*. To learn more about AWS Regions, see [Managing AWS Regions](#) in the *Amazon Web Services General Reference*.

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

### AWS Management Console

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

### AWS command line tools

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

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

### AWS SDKs

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

### Amazon Macie REST API

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

## Pricing for Amazon Macie

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



Macie pricing is based on two dimensions—evaluating and monitoring S3 buckets for security and access control, and analyzing S3 objects to discover and report sensitive data in those objects. To help you understand and forecast the cost of using Macie, Macie provides estimated usage costs for your account. You can [view these estimates \(p. 193\)](#) on the Amazon Macie console and access them with the Amazon Macie API.

Depending on how you use the service, you might incur additional costs for using other AWS services in combination with certain Macie features, such as retrieving bucket data from Amazon S3 and using customer managed AWS KMS keys to decrypt objects for analysis. For more information, see [Amazon Macie pricing](#).

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

To help you understand and forecast the cost of using Macie after the free trial ends, Macie provides you with estimated usage costs based on your use of Macie during the trial. Your usage data also indicates the amount of time that remains before your free trial ends. You can [view this data \(p. 198\)](#) on the Amazon Macie console and access it with the Amazon Macie API.

## Related services

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

### **AWS Security Hub**

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

To learn more about Security Hub, see the [AWS Security Hub User Guide](#). To learn about using Macie and Security Hub together, see [Amazon Macie integration with AWS Security Hub \(p. 166\)](#).

### **Amazon GuardDuty**

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

To learn more about GuardDuty, see the [Amazon GuardDuty User Guide](#).

To learn about additional AWS security services, see [Security, Identity, and Compliance on AWS](#).

# Getting started with Amazon Macie

This tutorial provides a hands-on introduction to Amazon Macie.

## Tasks

- [Before you begin](#) (p. 5)
- [Step 1: Enable Amazon Macie](#) (p. 5)
- [Step 2: Configure a repository for sensitive data discovery results](#) (p. 6)
- [Step 3: Create a job to discover sensitive data](#) (p. 6)
- [Step 4: Review your findings](#) (p. 7)

## Before you begin

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

## Step 1: Enable Amazon Macie

After you set up the required permissions, you can enable Macie. Follow these steps to enable Macie.

### To enable Macie

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. By using the AWS Region selector in the upper-right corner of the page, select the Region in which you want to enable Macie.
3. Choose **Get started**.
4. (Optional) When you enable Macie, Macie creates a service-linked role that grants Macie the permissions that it requires to call other AWS services on your behalf. To learn more about this role, see [Service-linked roles for Amazon Macie](#) (p. 205).
5. Choose **Enable Macie**.

Within minutes, Macie generates an inventory of the Amazon Simple Storage Service (Amazon S3) buckets for your account in the current Region. Macie also begins monitoring the buckets for security and access control.

To review your bucket inventory, choose **S3 buckets** in the navigation pane on the console. To then display details about a bucket, choose the bucket's name in the table. The details panel displays statistics and other information that provides insight into the security and privacy of the bucket's data. To learn more about these details, see [Reviewing your S3 bucket inventory](#) (p. 18).

## Step 2: Configure a repository for sensitive data discovery results

With Macie, you detect sensitive data by creating and running sensitive data discovery jobs. A sensitive data discovery job analyzes objects in S3 buckets to determine whether the objects contain sensitive data. If Macie discovers sensitive data in an object, Macie creates a *sensitive data finding*. A *sensitive data finding* is a detailed report of sensitive data that Macie found in an object.

Macie also creates a *sensitive data discovery result* for each object that you configure a job to analyze. A *sensitive data discovery result* is a record that logs details about the analysis of an object. This includes objects that don't contain sensitive data, and therefore don't produce a sensitive data finding, and objects that Macie can't analyze due to issues such as permissions settings. If an object does contain sensitive data, the sensitive data discovery result includes data from the corresponding sensitive data finding. It provides additional information too.

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

To learn how to configure a repository for your discovery results, see [Storing and retaining sensitive data discovery results \(p. 96\)](#).

## Step 3: Create a job to discover sensitive data

In Macie, sensitive data discovery jobs analyze objects in S3 buckets to detect and report sensitive data in those objects. Each job can use the built-in, managed data identifiers that Macie provides and custom data identifiers that you create. For information about the types of data that Macie can analyze, see [Discovering sensitive data \(p. 36\)](#). For information about the types of sensitive data that Macie can detect, see [Using managed data identifiers \(p. 37\)](#).

Follow these steps to create a job that runs once, immediately after you create it, and uses default settings. To learn how to create a job that runs periodically or uses custom settings, see [Creating a sensitive data discovery job \(p. 65\)](#).

### To create a sensitive data discovery job

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Jobs**.
3. Choose **Create job**.
4. For the **Choose S3 buckets** step, choose **Select specific buckets**.

Macie displays a complete inventory of the S3 buckets for your account in the current Region.

5. Select the check box for each bucket that you want the job to analyze.

#### Tip

To find specific buckets more easily, you can enter filter criteria in the filter bar above the table. You can also sort the inventory by choosing a column heading in the table.

6. When you finish selecting buckets, choose **Next**.
7. For the **Review S3 buckets** step, review and verify your bucket selections. Then choose **Next**.
8. For the **Refine the scope** step, choose **One-time job**, and then choose **Next**.
9. For the **Select managed data identifiers** step, choose **All**, and then choose **Next**.

10. For the **Select custom data identifiers** step, choose **Next**.
11. For the **Enter a name and description** step, enter a name and, optionally, a description of the job. Then choose **Next**.
12. For the **Review and create** step, review the job's configuration settings and verify that they're correct.

You can also review the total estimated cost (in US Dollars) of running the job. To learn more about this estimate, see [Forecasting the cost of a sensitive data discovery job \(p. 87\)](#).

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

Macie immediately starts running the job. You can then [monitor and check the status of the job \(p. 85\)](#).

## Step 4: Review your findings

Macie automatically monitors your S3 buckets for security and access control, and it creates policy findings to report any potential policy violations. If you create and run a sensitive data discovery job, Macie creates sensitive data findings to report any sensitive data that it discovers. To learn about findings, see [Analyzing findings \(p. 104\)](#).

Follow these steps to review the details of your findings.

### To review your findings

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Findings**.
3. (Optional) To filter the findings by specific criteria, enter the criteria in the filter bar above the table.
4. To view the details of a specific finding, choose any field other than the check box for the finding. The details panel displays information for the finding.

# Monitoring Amazon S3 data with Amazon Macie

When you enable Amazon Macie for your AWS account, Macie automatically generates and begins maintaining a complete inventory of your Amazon Simple Storage Service (Amazon S3) buckets in the current AWS Region. Macie also begins monitoring and evaluating the buckets for security and access control. If Macie detects an event that reduces the security or privacy of an S3 bucket, Macie creates a [policy finding \(p. 105\)](#) for you to review and remediate as necessary.

To also monitor S3 buckets for the presence of sensitive data, you can create and run [sensitive data discovery jobs \(p. 56\)](#) that analyze bucket objects on a daily, weekly, or monthly basis. If you do this and Macie detects sensitive data in an object, Macie creates a [sensitive data finding \(p. 106\)](#) to notify you of the sensitive data that Macie found.

In addition to findings, Macie provides constant visibility into the security and privacy of your Amazon S3 data. To assess the security posture of your data and determine where to take action, you can use the **Summary** dashboard on the console. This dashboard provides a snapshot of aggregated statistics for your Amazon S3 data. The statistics include data for key security metrics such as the number of buckets that are publicly accessible, don't encrypt new objects by default, or are shared with other AWS accounts. The dashboard also displays groups of aggregated findings data for your account—for example, the names of 1–5 buckets that have the most findings for the preceding seven days. You can drill down on each statistic to view its supporting data. If you prefer to query the statistics programmatically, you can use the [Amazon S3 Data Source Statistics](#) resource of the Amazon Macie API.

For deeper analysis and evaluation, Macie also provides detailed information and statistics for individual buckets in your inventory. This includes breakdowns of each bucket's public access and encryption settings, and the size and number of objects that Macie can analyze to detect sensitive data in the bucket. The inventory also indicates whether any sensitive data discovery jobs are configured to analyze objects in a bucket and, if so, when one of those jobs most recently ran. You can browse, sort, and filter the inventory by using the Amazon Macie console or the [Amazon S3 Data Source](#) resource of the Amazon Macie API.

If you're the Macie administrator for an organization, you can access statistical and other data for S3 buckets that are owned by member accounts in your organization. You can also access policy findings that Macie creates for the buckets, and create sensitive data discovery jobs to detect sensitive data in the buckets. This means that you can use Macie to evaluate and monitor your organization's security posture across your Amazon S3 environment. For more information, see [Managing multiple accounts \(p. 184\)](#).

## Topics

- [How Amazon Macie monitors Amazon S3 data \(p. 8\)](#)
- [Assessing your Amazon S3 security posture with Amazon Macie \(p. 12\)](#)
- [Analyzing your Amazon S3 security posture with Amazon Macie \(p. 17\)](#)
- [Allowing Amazon Macie to access S3 buckets and objects \(p. 32\)](#)

## How Amazon Macie monitors Amazon S3 data

When you enable Amazon Macie for your AWS account, Macie creates an AWS Identity and Access Management (IAM) [service-linked role \(p. 205\)](#) for your account in the current AWS Region. The

permissions policy for this role allows Macie to monitor AWS resources and call other AWS services on your behalf. By using this role, Macie generates and maintains a complete inventory of your Amazon Simple Storage Service (Amazon S3) buckets in the Region, and Macie monitors and evaluates the buckets for security and access control.

If you're the Macie administrator for an organization, the inventory includes statistical and other data about S3 buckets that are owned by your account and by member accounts in your organization. With this data, you can use Macie to monitor and evaluate your organization's security posture across your Amazon S3 environment. For more information, see [Managing multiple accounts \(p. 184\)](#).

### Topics

- [Key components \(p. 9\)](#)
- [Data refreshes \(p. 10\)](#)
- [Additional considerations \(p. 11\)](#)

## Key components

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

### Gathering metadata and calculating statistics

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

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

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

### Monitoring bucket security and privacy

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

For example, if you enable default encryption for a bucket, Macie updates all data about the bucket's default encryption settings. Similarly, if you update an existing bucket policy or add a policy to a bucket, Macie analyzes the policy and updates the relevant data in your inventory.

Macie monitors and analyzes data for the following CloudTrail events:

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

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

#### Tip

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

### Evaluating bucket security and access control

To evaluate bucket-level security and access control, Macie uses automated, logic-based reasoning to analyze resource-based policies that apply to a bucket. Macie also analyzes the account- and bucket-level permissions settings that apply to a bucket. This analysis factors bucket policies, bucket-level ACLs, and block public access settings for the account and the bucket.

For resource-based policies, Macie uses [Zelkova](#). Zelkova is an automated reasoning engine that translates IAM policies into logical statements and runs a suite of general-purpose and specialized logical solvers (*satisfiability modulo theories*) against the decision problem. Macie applies Zelkova repeatedly to a policy with increasingly specific queries to characterize the classes of behaviors that the policy allows. To learn more about the nature of the solvers that Zelkova uses, see [Satisfiability Modulo Theories](#).

#### Important


To perform the preceding tasks for a bucket, Macie must be allowed to access the bucket. If a bucket's permissions settings prevent Macie from retrieving metadata for the bucket or the bucket's objects, Macie can only provide a subset of information about the bucket, such as the bucket's name and creation date. Macie can't perform any additional tasks for the bucket. For more information, see [Allowing Macie to access S3 buckets and objects](#) (p. 32).

## Data refreshes

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


To determine when Macie most recently retrieved both bucket and object metadata for your account as part of the daily refresh cycle, you can refer to the **Last updated** field on the console. This field appears on the **Summary** dashboard and the **S3 buckets** page, and in the [bucket details panel](#) (p. 20). (If you use the Amazon Macie API to query inventory data, the `lastUpdated` field provides this information.) If you're the Macie administrator for an organization, the **Last updated** field indicates the earliest date and time when Macie retrieved the data for an account in your organization.

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

- You refresh your inventory data by choosing refresh () on the Amazon Macie console. You can refresh the data as frequently as every five minutes.
- You send a [DescribeBuckets](#) request to the Amazon Macie API programmatically, and you haven't sent a **DescribeBuckets** request within the preceding five minutes.

- Macie detects a relevant AWS CloudTrail event.

Macie can also retrieve the latest object metadata for a specific bucket if you choose to manually refresh that data. This can be helpful if you recently created a bucket or made significant changes to a bucket's objects during the past 24 hours. To manually refresh object metadata for a bucket, choose refresh

 in the **Object statistics** section of the [bucket details panel](#) (p. 20) on the console. This feature is available for buckets that contain 30,000 or fewer objects.

Each time Macie retrieves bucket or object metadata, Macie automatically updates all the relevant data in your inventory. If Macie detects differences that affect the security or privacy of a bucket, Macie immediately begins evaluating and analyzing the changes. When the analysis is complete, Macie updates the relevant data in your inventory. If any differences reduce the security or privacy of a bucket, Macie also creates the appropriate [policy findings](#) (p. 105) for you to review and remediate as necessary.


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

## Additional considerations

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

- Inventory data applies only to S3 buckets in the current AWS Region. To access the data for additional Regions, enable and use Macie in each additional Region.
- If you're the Macie administrator for an organization, you can access inventory data for a member account only if Macie is enabled for that account in the current Region.
- If a bucket's permissions settings prevent Macie from retrieving information about the bucket or the bucket's objects, Macie can't evaluate and monitor the security and privacy of the bucket's data or provide detailed information about the bucket.

To help you identify a bucket where this is the case, Macie does the following:

- In your bucket inventory, Macie displays a warning icon () for the bucket. For the bucket's details, Macie displays only a subset of fields and data: the account ID for the AWS account that owns the bucket; the bucket's name, Amazon Resource Name (ARN), creation date, and Region; and, the date and time when Macie most recently retrieved both bucket and object metadata for the bucket as part of the daily refresh cycle. If you use the Amazon Macie API to query inventory data, Macie provides an error code and message for the bucket and the value for most of the bucket's properties is null.
- On the **Summary** dashboard, the bucket has a value of *Unknown* for the **Public access**, **Encryption**, and **Sharing** statistics. (If you use the Amazon Macie API to query the statistics, the bucket has a value of *unknown* for these statistics.) In addition, Macie excludes the bucket when it calculates data for **Storage** and **Objects** statistics.

To investigate the issue, review the bucket's policy and permissions settings in Amazon S3. For example, the bucket might have a restrictive bucket policy. For more information, see [Allowing Macie to access S3 buckets and objects](#) (p. 32).

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

To monitor object-level operations and identify potential security risks, we recommend that you use the Amazon S3 protection feature of Amazon GuardDuty. This feature monitors object-level, Amazon



S3 data events and analyzes them for malicious and suspicious activity. For more information, see [Amazon S3 protection in Amazon GuardDuty](#) in the *Amazon GuardDuty User Guide*.

- If you manually refresh object metadata for a specific bucket, Macie temporarily reports *Unknown* for encryption statistics that apply to the objects. The next time Macie performs the daily data refresh (within 24 hours), Macie re-evaluates the encryption metadata for the objects and reports quantitative data for the statistics again.
- In rare cases, Macie might not be able to determine whether a bucket is publicly accessible, is shared with another AWS account, or requires server-side encryption of new objects. For example, a temporary issue might prevent Macie from retrieving and analyzing the requisite data. Or Macie might not be able to fully determine whether one or more policy statements grant access to an external entity. In these cases, Macie reports *Unknown* for the relevant statistics and fields in the inventory. To investigate these cases, review the bucket's policy and permissions settings in Amazon S3.

Also note that Macie generates policy findings only if the security or privacy of a bucket is reduced after you enable Macie for your account. For example, if you disable default encryption for a bucket after you enable Macie, Macie generates a **Policy:IAMUser/S3BucketEncryptionDisabled** finding for the bucket. However, if default encryption was disabled for a bucket when you enabled Macie and default encryption continues to be disabled, Macie doesn't generate a **Policy:IAMUser/S3BucketEncryptionDisabled** finding for the bucket.

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

## Assessing your Amazon S3 security posture with Amazon Macie

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

The **Summary** dashboard provides a snapshot of aggregated statistics for your Amazon S3 data in the current AWS Region. The statistics include data for key security metrics such as the number of buckets that are publicly accessible or don't encrypt new objects by default. The dashboard also displays groups of aggregated findings data for your account—for example, the types of findings that had the highest number of occurrences during the preceding seven days. If you're the Macie administrator for an organization, the dashboard includes aggregated statistics and data for member accounts in your organization.

To perform deeper analysis, you can drill down and review the supporting data for individual items on the dashboard. You can also [review and analyze your S3 bucket inventory](#) (p. 17) by using the Amazon Macie console, or query and analyze inventory data by using the [Amazon S3 Data Source](#) resource of the Amazon Macie API.

### Topics

- [Displaying the Summary dashboard](#) (p. 13)
- [Understanding components of the Summary dashboard](#) (p. 13)
- [Understanding S3 bucket statistics on the Summary dashboard](#) (p. 15)

## Displaying the Summary dashboard

On the Amazon Macie console, the **Summary** dashboard provides a snapshot of aggregated statistics and findings data for your Amazon S3 data in the current AWS Region. If you prefer to query and review this data programmatically, you can use the [Amazon S3 Data Source Statistics](#) resource of the Amazon Macie API.

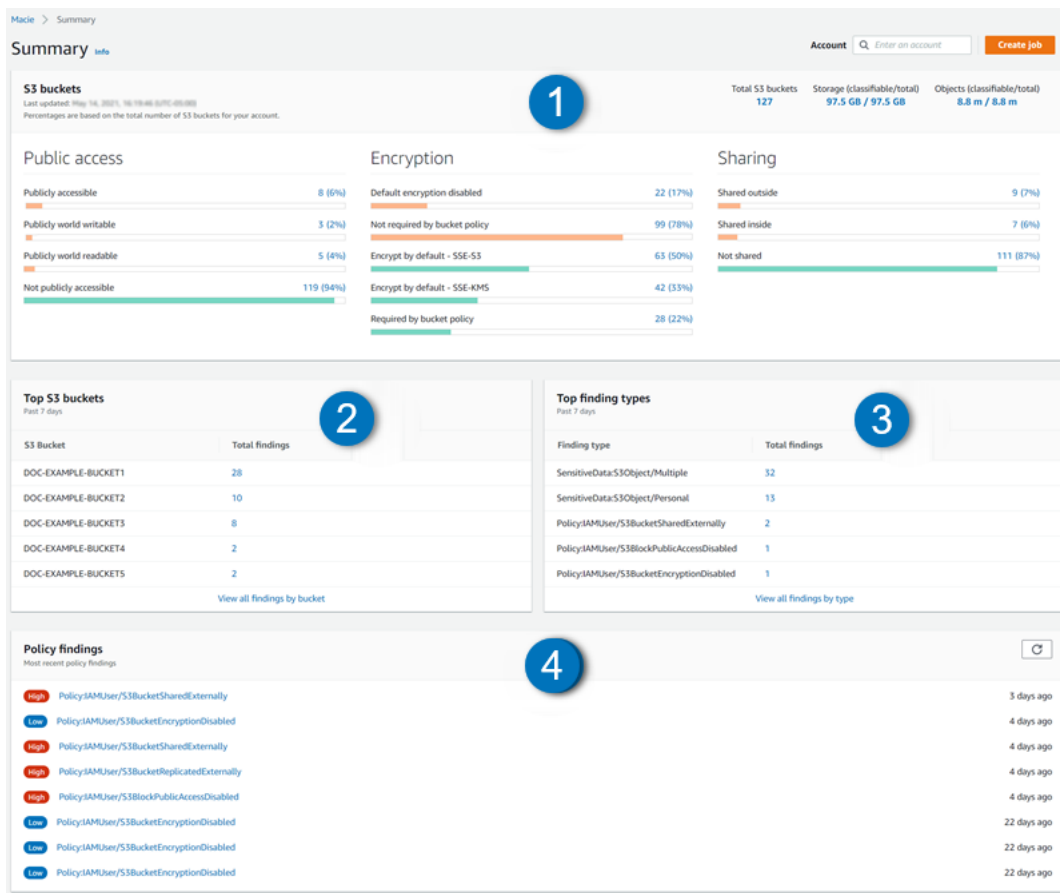
### To display the Summary dashboard

1. Open the Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Summary**. Macie displays the **Summary** dashboard.
3. To determine when Macie most recently retrieved both bucket and object metadata for your account, refer to the **Last updated** field at the top of the dashboard. For more information, see [Data refreshes \(p. 10\)](#).
4. To review the supporting data for an item on the dashboard, choose the item.

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

## Understanding components of the Summary dashboard

On the **Summary** dashboard, statistics and data are organized into four sections, as shown in the following image.



Each section provides insight into key metrics or recent findings data that can help you assess the security and privacy of your Amazon S3 data in the current AWS Region.

## 1. S3 buckets

This section provides statistics about the amount of data that you store in Amazon S3 and how much of that data Macie can analyze to detect sensitive data. It also provides statistics that indicate potential security and privacy risks for the data. For details about each statistic, see [Understanding S3 bucket statistics on the dashboard](#) (p. 15).

This section also indicates when Macie most recently retrieved bucket and object metadata from Amazon S3 as part of the daily refresh cycle. You can find this information in the **Last updated** field. For more information, see [Data refreshes](#) (p. 10).

## 2. Top S3 buckets

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

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

This section is empty if Macie didn't create any findings during the preceding seven days.

### 3. Top finding types

This section lists the [types of findings](#) (p. 105) that had the highest number of occurrences during the preceding seven days, for as many as five types of findings. It also indicates the number of findings that Macie created for each type.

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

This section is empty if Macie didn't create any findings during the preceding seven days.

### 4. Policy findings

This section lists the [policy findings](#) (p. 105) that Macie created or updated most recently, for as many as ten findings. To display the details of a particular finding, choose the finding.

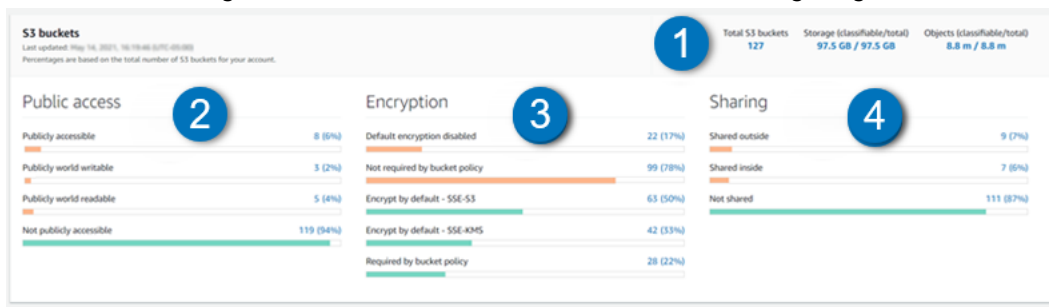
This section is empty if Macie didn't create or update any policy findings during the preceding seven days.

Note that findings data on the **Summary** dashboard doesn't include findings that were suppressed by a [suppression rule](#) (p. 152).

## Understanding S3 bucket statistics on the Summary dashboard

The **S3 buckets** section of the **Summary** dashboard provides statistics about the amount of data that you store in Amazon S3 in the current AWS Region and how much of that data Macie can analyze to detect sensitive data. It also provides statistics that can help you identify and investigate potential security risks. For example, you might use this data to identify S3 buckets that are publicly accessible or don't encrypt new objects by default, and then [create a sensitive data discovery job](#) (p. 65) to determine whether the buckets also contain sensitive data.

The statistics are organized into four sections, as shown in the following image.



### 1. Storage and sensitive data discovery

The statistics at the top of the **S3 buckets** section indicate how much data you store in Amazon S3 and how much of that data Macie can analyze to detect sensitive data:

- **Total S3 buckets** – The total number of S3 buckets, including buckets that don't contain any objects.
- **Storage**
  - **Classifiable** – The total storage size of all the objects that Macie can analyze in the buckets.
  - **Total** – The total storage size of all the objects in the buckets, including objects that Macie can't analyze.

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

- **Objects**

- **Classifiable** – The total number of objects that Macie can analyze in the buckets.
- **Total** – The total number of objects in the buckets, including objects that Macie can't analyze.

In the preceding statistics, data and objects are *classifiable* if they use a supported Amazon S3 storage class (S3 Intelligent-Tiering, S3 One Zone-IA, S3 Standard, or S3 Standard-IA) and have a file name extension for a [supported file or storage format \(p. 89\)](#). You can detect sensitive data in these objects by creating and running sensitive data discovery jobs.

Note that the **Storage** and **Objects** statistics don't include data about objects in buckets that Macie isn't allowed to access. To identify buckets where this is the case, you can [review your bucket inventory \(p. 18\)](#). If the warning icon (⚠) appears next to a bucket's name in your inventory, Macie isn't allowed to access the bucket.

## 2. Public access

This section indicates how many S3 buckets are or aren't publicly accessible:

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

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

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

In certain cases, this section also displays values for *Unknown*. If these values appear, Macie wasn't able to evaluate the public access settings for the specified number and percentage of buckets. For example, a temporary issue or the buckets' permissions settings prevented Macie from retrieving the requisite data. Or Macie wasn't able to fully determine whether one or more policy statements allow an external entity to access the buckets.

## 3. Encryption

This section indicates how many S3 buckets are or aren't configured to encrypt new objects automatically, and how many S3 buckets do or don't require server-side encryption of objects when objects are uploaded to the buckets:

- **Default encryption disabled** – The number and percentage of buckets that don't encrypt new objects automatically. Default encryption is disabled for these buckets.
- **Not required by bucket policy** – The number and percentage of buckets whose bucket policies don't require server-side encryption of new objects. For these buckets, [PutObject](#) requests don't have to specify a valid, server-side encryption option.
- **Encrypt by default – SSE-S3** – The number and percentage of buckets that encrypt new objects automatically using an Amazon S3 managed key. Default encryption is enabled for these buckets.

- **Encrypt by default – SSE-KMS** – The number and percentage of buckets that encrypt new objects automatically using an AWS KMS key. Default encryption is enabled for these buckets.
- **Required by bucket policy** – The number and percentage of buckets whose bucket policies require server-side encryption of new objects. For these buckets, **PutObject** requests must specify a valid, server-side encryption option. Otherwise, Amazon S3 denies the request.

Note that the totals in this section might exceed the total number of buckets in your inventory. This is because a bucket can have a combination of encryption settings. For example, a bucket might not be configured to encrypt new objects automatically and it might not have a bucket policy that requires server-side encryption of new objects.

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

To determine the values in this section, Macie analyzes the default encryption settings and, if applicable, the bucket policy for each bucket. For information about default encryption settings, see [Setting default server-side encryption behavior for Amazon S3 buckets](#) in the *Amazon Simple Storage Service User Guide*. For information about using bucket policies to require server-side encryption of new objects, see [How to prevent uploads of unencrypted objects to Amazon S3](#) on the *AWS Security Blog*.

In certain cases, this section also displays values for *Unknown*. If these values appear, Macie wasn't able to evaluate the default encryption settings or bucket policy for the specified number and percentage of buckets. For example, a temporary issue or the buckets' permissions settings prevented Macie from retrieving the requisite data. Or Macie wasn't able to fully determine whether the buckets' policies require server-side encryption of new objects.

#### 4. Sharing

This section indicates how many S3 buckets are or aren't shared with other AWS accounts:

- **Shared outside** – The number and percentage of buckets that are shared with accounts that aren't in the same organization.
- **Shared inside** – The number and percentage of buckets that are shared with accounts in the same organization.
- **Not shared** – The number and percentage of buckets that aren't shared with other accounts.

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

To determine the values in this section, Macie analyzes the bucket policy and ACL for each bucket. In addition, an *organization* is defined as a set of Macie accounts that are centrally managed as a group of related accounts through AWS Organizations or by Macie invitation.

In certain cases, this section also displays values for *Unknown*. If these values appear, Macie wasn't able to determine whether the specified number and percentage of buckets are shared with another account. For example, a temporary issue or the buckets' permissions settings prevented Macie from retrieving the requisite data. Or Macie wasn't able to fully determine whether the buckets' policies or ACLs are configured to share the buckets with another account.

## Analyzing your Amazon S3 security posture with Amazon Macie

To help you perform in-depth analysis and evaluate the security posture of your Amazon Simple Storage Service (Amazon S3) data, Amazon Macie maintains a complete inventory of your S3 buckets in each AWS Region where you use Macie. To learn how Macie maintains this inventory for you, see [How Macie](#)

[monitors Amazon S3 data \(p. 8\)](#). If you're the Macie administrator for an organization, the inventory includes data for S3 buckets that are owned by member accounts in your organization.

By using this inventory, you can review your Amazon S3 data estate, and examine details and statistics for key security settings and metrics that apply to individual S3 buckets. For example, you can access breakdowns of each bucket's public access and encryption settings, and the size and number of objects that Macie can analyze to detect sensitive data in each bucket. You can also determine whether you've configured any sensitive data discovery jobs to analyze data in a bucket. If you have, the inventory indicates when one of those jobs most recently ran.

You can browse, sort, and filter inventory data by using the **S3 buckets** page on the Amazon Macie console. You can also access your inventory data programmatically by using the [Amazon S3 Data Source](#) resource of the Amazon Macie API.

#### Topics

- [Reviewing your S3 bucket inventory with Amazon Macie \(p. 18\)](#)
- [Filtering your S3 bucket inventory with Amazon Macie \(p. 24\)](#)

## Reviewing your S3 bucket inventory with Amazon Macie

On the Amazon Macie console, the **S3 buckets** page provides detailed insight into the security and privacy of your Amazon Simple Storage Service (Amazon S3) data. With this page, you can review and analyze a complete inventory of your S3 buckets in the current AWS Region, and review detailed information and statistics for individual buckets. If you're the Macie administrator for an organization, your inventory includes details and statistics for S3 buckets that are owned by member accounts in your organization.

The **S3 buckets** page also indicates when Macie most recently retrieved both bucket and object metadata for your account as part of the daily refresh cycle. You can find this information in the **Last updated** field at the top of the page. For more information, see [Data refreshes \(p. 10\)](#).

Note that most inventory data is limited to the buckets that Macie is allowed to access for your account. If a bucket's permissions settings prevent Macie from retrieving information about the bucket or the bucket's objects, Macie can only provide a subset of information about the bucket. If this is the case for a particular bucket, Macie displays a warning icon (⚠) and message for the bucket in your bucket inventory. For the bucket's details, Macie displays only a subset of fields and data: the account ID for AWS account that owns the bucket; the bucket's name, Amazon Resource Name (ARN), creation date, and Region; and, the date and time when Macie most recently retrieved both bucket and object metadata for the bucket as part of the daily refresh cycle. To investigate the issue, review the bucket's policy and permissions settings in Amazon S3. For example, the bucket might have a restrictive bucket policy. For more information, see [Allowing Macie to access S3 buckets and objects \(p. 32\)](#).

If you prefer to access and query your inventory data programmatically, you can use the [Amazon S3 Data Source](#) resource of the Amazon Macie API.

#### Topics

- [Viewing your S3 bucket inventory \(p. 18\)](#)
- [Viewing the details of S3 buckets \(p. 20\)](#)


## Viewing your S3 bucket inventory


The **S3 buckets** page on the Amazon Macie console provides information about your S3 buckets in the current AWS Region. On this page, a table displays summary information for each bucket in your inventory. To customize your view, you can sort and filter the table.



If you choose a bucket in the table, the details panel displays additional information about the bucket. This includes details and statistics for settings and metrics that provide insight into the security and privacy of the bucket's data.

### To view your S3 bucket inventory

1. Open the Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **S3 buckets**. The **S3 buckets** page opens and displays the number of buckets in your inventory and a table of the buckets.
3. At the top of the page, optionally choose refresh () to retrieve the latest bucket metadata from Amazon S3.

If the information icon () appears next to any bucket names, we recommend that you do this. This icon indicates that a bucket was created during the past 24 hours, possibly after Macie last retrieved bucket and object metadata from Amazon S3 as part of the [daily refresh cycle \(p. 10\)](#).

4. On the **S3 buckets** page, use the table to review a subset of information about each bucket in your inventory:
  - **Bucket** – The name of the bucket.
  - **Account** – The account ID of the AWS account that owns the bucket.
  - **Classifiable objects** – The total number of objects that Macie can analyze to detect sensitive data in the bucket.
  - **Classifiable size** – The total storage size of all the objects that Macie can analyze to detect sensitive data in the bucket.

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

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

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



- **Latest job run** – If any one-time or periodic sensitive data discovery jobs are configured to analyze objects in the bucket, the value for this field indicates the most recent time when one of those jobs started to run. Otherwise, this field is empty.

In the preceding data, objects are *classifiable* if they use a supported Amazon S3 storage class and have a file name extension for a supported file or storage format. You can detect sensitive data in these objects by creating and running a sensitive data discovery job: select the check box for each bucket that contains objects to analyze, and then choose **Create job**. For more information, see [Discovering sensitive data \(p. 36\)](#).

5. To analyze your inventory by using the table, do any of the following:
  - To sort the table by a specific field, click the column heading for the field. To change the sort order, click the column heading again.
  - To filter the table and display only those buckets that have a specific value for a field, place your cursor in the filter bar, and then add a filter condition for the field. To further refine the results, add filter conditions for additional fields. For more information, see [Filtering your S3 bucket inventory \(p. 24\)](#).
6. To review details and statistics for a particular bucket, choose the bucket's name in the table, and then refer to the details panel.



**Tip**


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


## Viewing the details of S3 buckets

On the Amazon Macie console, you can use the details panel on the **S3 buckets** page to review statistics and other information about individual S3 buckets in your bucket inventory. This includes details and statistics for settings and metrics that provide insight into the security and privacy of a bucket's data.

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

### To view the details of an S3 bucket



1. Open the Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **S3 buckets**.
3. On the **S3 buckets** page, optionally choose refresh () to retrieve the latest bucket metadata from Amazon S3.

If the information icon () appears next to any bucket names, we recommend that you do this. This icon indicates that a bucket was created during the past 24 hours, possibly after Macie last retrieved bucket and object metadata from Amazon S3 as part of the [daily refresh cycle \(p. 10\)](#).

4. In the **S3 buckets** table, choose the name of the bucket whose details you want to review. The details panel displays statistics and other information about the bucket.

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

- [Overview \(p. 20\)](#)
- [Object statistics \(p. 21\)](#)
- [Server-side encryption \(p. 22\)](#)
- [Sensitive data discovery \(p. 23\)](#)
- [Public access \(p. 23\)](#)
- [Replication \(p. 23\)](#)
- [Tags \(p. 24\)](#)

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

### Overview



This section provides general information about the bucket, such as the bucket's name, when the bucket was created, and the account ID of the AWS account that owns the bucket. The **Last updated** field

indicates when Macie most recently retrieved metadata from Amazon S3 for both the bucket and the bucket's objects as part of the [daily refresh cycle \(p. 10\)](#).

Of special note, the **Shared access** field indicates whether the bucket is shared with other AWS accounts and, if so, whether those accounts are internal to (part of) or external to (not part of) your organization. An *organization* is a set of Macie accounts that are centrally managed as a group of related accounts through AWS Organizations or by Macie invitation. To determine the value for this field, Macie analyzes the bucket policy and access control list (ACL) for the bucket. Note that this data is limited to bucket-level settings. It doesn't reflect any object-level settings for sharing specific objects with another account.

## Object statistics

This section provides information about the objects in the bucket, starting with the total number of objects in the bucket, the total storage size of all those objects, and the total storage size of all the objects that are compressed (.gz, .gzip, .zip) files. If versioning is enabled for the bucket, the size values are based on the size of the latest version of each object in the bucket.

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

### Note

If you refresh object metadata for a bucket, Macie temporarily reports *Unknown* for encryption statistics that apply to the objects. Macie will re-evaluate and update the data for these statistics when it performs the next daily refresh of bucket and object metadata, which is within 24 hours.

Additional statistics in this section can help you assess how much data Macie can analyze to detect sensitive data in the bucket.

### Classifiable objects

This section indicates the total number objects that Macie can analyze to detect sensitive data and the total storage size of those objects. These objects use a supported Amazon S3 storage class (S3 Intelligent-Tiering, S3 One Zone-IA, S3 Standard, or S3 Standard-IA) and have a file name extension for a supported file or storage format. This means that you can detect sensitive data in the objects by creating and running a sensitive data discovery job. For more information, see [Discovering sensitive data \(p. 36\)](#).

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

### Unclassifiable objects

This section indicates the total number of objects that Macie can't analyze to detect sensitive data and the total storage size of those objects. These objects don't use a supported Amazon S3 storage class (S3 Intelligent-Tiering, S3 One Zone-IA, S3 Standard, or S3 Standard-IA) or don't have a file name extension for a [supported file or storage format \(p. 89\)](#).

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

### Unclassifiable objects: Storage class

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

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

#### Unclassifiable objects: File type

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

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

#### Objects by encryption type

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

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

For information about each supported encryption type, see [Protecting data using encryption](#) in the *Amazon Simple Storage Service User Guide*.

## Server-side encryption

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

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

- **No** – The bucket doesn't have a bucket policy or the bucket's policy doesn't require server-side encryption of new objects. If a bucket policy exists, it doesn't require [PutObject](#) requests to include the `x-amz-server-side-encryption` header and it doesn't require the value for that header to be `AES256` or `aws:kms`.
- **Yes** – The bucket policy requires server-side encryption of new objects. This means that **PutObject** requests for the bucket must include the `x-amz-server-side-encryption` header and the value for that header must be `AES256` or `aws:kms`. Otherwise, Amazon S3 denies the request.
- **Unknown** – Macie wasn't able to evaluate the bucket policy to determine whether it requires server-side encryption of new objects.

For information about using bucket policies to require server-side encryption of new objects, see [How to prevent uploads of unencrypted objects to Amazon S3](#) on the *AWS Security Blog*.

The **Default encryption** field indicates whether default encryption is enabled for the bucket and, if so, the type of server-side encryption that's used:

- **AES256** – New objects are encrypted automatically with an Amazon S3 managed key. Default encryption is enabled for the bucket and it uses SSE-S3 encryption.
- **aws:kms** – New objects are encrypted automatically with an AWS KMS key, either an AWS managed KMS key or a customer managed KMS key. Default encryption is enabled for the bucket and it uses SSE-KMS encryption. The **KMS master key** field shows the Amazon Resource Name (ARN) or unique identifier (key ID) for the KMS key that's used.
- **None** – New objects aren't encrypted automatically. Default encryption is disabled for the bucket.

For information about configuring default encryption settings, see [Setting default server-side encryption behavior for Amazon S3 buckets](#) in the *Amazon Simple Storage Service User Guide*.

## Sensitive data discovery

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

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

### Tip

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

## Public access

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

- **Not public** – The bucket isn't publicly accessible.
- **Public** – The bucket is publicly accessible.
- **Unknown** – Macie wasn't able to evaluate all the public access settings for the bucket.

Note that this data is limited to account- and bucket-level settings. It doesn't reflect object-level settings that enable public access to specific objects in a bucket.

To learn about Amazon S3 settings for managing public access to buckets and bucket data, see [Identity and access management in Amazon S3](#) and [Blocking public access to your Amazon S3 storage](#) in the *Amazon Simple Storage Service User Guide*.

## Replication

In this section, the **Replicated** field indicates whether the bucket is configured to replicate objects to buckets that are owned by other AWS accounts. If the bucket is configured to do this, this section also lists the account IDs for those accounts.

The **Replicated externally** field indicates whether bucket objects are replicated to AWS accounts that are external to (aren't part of) your organization. An *organization* is a set of Macie accounts that are centrally managed as a group of related accounts through AWS Organizations or by Macie invitation.

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

## Tags

If one or more tags are associated with the bucket, this section appears in the panel and it lists those tags. Tags are custom labels that you can associate with specific types of AWS resources, including S3 buckets. Each tag consists of a required tag key and an optional tag value.

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

## Filtering your S3 bucket inventory with Amazon Macie

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

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

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

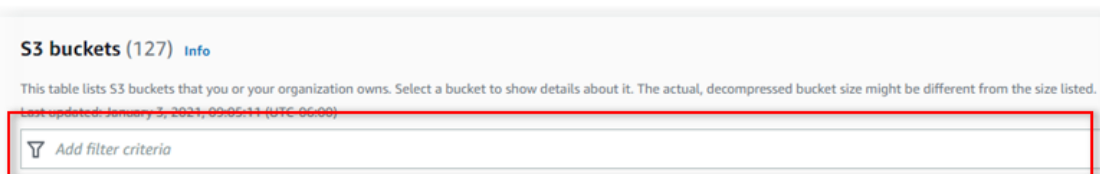
How you define and apply filter conditions depends on whether you use the Amazon Macie console or the Amazon Macie API.

### Topics

- [Filtering your inventory on the Amazon Macie console](#) (p. 24)
- [Filtering your inventory programmatically with the Amazon Macie API](#) (p. 26)

## Filtering your inventory on the Amazon Macie console

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



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

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

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

After you add a condition, Macie applies the condition's criteria and adds the condition to a filter box in the filter bar, as shown in the following image.




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

As you add more conditions, Macie applies their criteria and adds them to the filter bar. If you add multiple conditions, Macie uses AND logic to join the conditions and evaluate the filter criteria. This means that a bucket meets the filter criteria only if it matches all the conditions in the filter.

You can refer to the filter bar at any time to see which criteria you've applied.

### To filter your inventory by using the console

1. Open the Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **S3 buckets**. The **S3 buckets** page opens and displays the number of buckets in your inventory and a table of the buckets.
3. To retrieve the latest bucket metadata from Amazon S3, choose refresh () at the top of the page.
4. Place your cursor in the filter bar, and then choose the field to use for the condition.
5. Choose or enter the appropriate type of value for the field, keeping the following tips in mind.

#### Dates, times, and time ranges

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

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

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

#### Numbers and numeric ranges

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

- To define a fixed numeric range, use the **From** and **To** boxes to specify the lowest and highest numbers in the range, respectively.

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

#### Text (string) values

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

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

6. When you finish adding a value for the field, choose **Apply**. Macie applies the filter criteria and adds the condition to a filter box in the filter bar.

#### Tip

For many fields, you can change a condition's operator from *equals* to *not equals* by choosing the equals icon (●) in the filter box. If you do this, Macie changes the operator to *not equals* and displays the not equals icon (⊄) in the filter box. To switch to the *equals* operator again, choose the not equals icon.

7. Repeat steps 4 through 6 for each additional condition that you want to add.
8. To remove a condition, choose the remove condition icon (⊗) in the filter box for the condition.
9. To change a condition, remove the condition by choosing the remove condition icon (⊗) in the filter box for the condition. Then repeat steps 4 through 6 to add a condition with the correct settings.

## Filtering your inventory programmatically with the Amazon Macie API

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

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

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

#### Examples

- [Example 1: Find buckets by bucket name \(p. 30\)](#)
- [Example 2: Find buckets that are publicly accessible \(p. 30\)](#)
- [Example 3: Find buckets that contain unencrypted objects \(p. 30\)](#)
- [Example 4: Find buckets that aren't monitored by a job \(p. 31\)](#)

- [Example 5: Find buckets that replicate data to external accounts \(p. 31\)](#)
- [Example 6: Find buckets based on multiple criteria \(p. 32\)](#)

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

### Example of a `buckets` array

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

```
{
  "buckets": [
    {
      "accountId": "123456789012",
      "bucketArn": "arn:aws:s3:::DOC-EXAMPLE-BUCKET1",
      "bucketCreatedAt": "2020-05-18T19:54:00+00:00",
      "bucketName": "DOC-EXAMPLE-BUCKET1",
      "allowsUnencryptedObjectUploads": "FALSE",
      "classifiableObjectCount": 13,
      "classifiableSizeInBytes": 1592088,
      "jobDetails": {
        "isDefinedInJob": "TRUE",
        "isMonitoredByJob": "TRUE",
        "lastJobId": "08c81dc4a2f3377fae45c9ddaexample",
        "lastJobRunTime": "2021-04-26T14:55:30.270000+00:00"
      },
      "lastUpdated": "2021-04-30T07:33:06.337000+00:00",
      "objectCount": 13,
      "objectCountByEncryptionType": {
        "customerManaged": 0,
        "kmsManaged": 2,
        "s3Managed": 7,
        "unencrypted": 4,
        "unknown": 0
      },
      "publicAccess": {
        "effectivePermission": "NOT_PUBLIC",
        "permissionConfiguration": {
          "accountLevelPermissions": {
            "blockPublicAccess": {
              "blockPublicAcls": true,
              "blockPublicPolicy": true,
              "ignorePublicAcls": true,
              "restrictPublicBuckets": true
            }
          },
          "bucketLevelPermissions": {
            "accessControlList": {
              "allowsPublicReadAccess": false,
              "allowsPublicWriteAccess": false
            },
            "blockPublicAccess": {
              "blockPublicAcls": true,
              "blockPublicPolicy": true,
              "ignorePublicAcls": true,
              "restrictPublicBuckets": true
            }
          },
          "bucketPolicy": {
            "allowsPublicReadAccess": false,
            "allowsPublicWriteAccess": false
          }
        }
      }
    }
  ]
}
```



```

    }
  },
  "region": "us-east-1",
  "replicationDetails": {
    "replicated": false,
    "replicatedExternally": false,
    "replicationAccounts": []
  },
  "serverSideEncryption": {
    "kmsMasterKeyId": null,
    "type": "NONE"
  },
  "sharedAccess": "NOT_SHARED",
  "sizeInBytes": 4549746,
  "sizeInBytesCompressed": 0,
  "tags": [
    {
      "key": "Division",
      "value": "HR"
    },
    {
      "key": "Team",
      "value": "Recruiting"
    }
  ],
  "unclassifiableObjectCount": {
    "fileType": 0,
    "storageClass": 0,
    "total": 0
  },
  "unclassifiableObjectSizeInBytes": {
    "fileType": 0,
    "storageClass": 0,
    "total": 0
  },
  "versioning": false
},
{
  "accountId": "123456789012",
  "bucketArn": "arn:aws:s3:::DOC-EXAMPLE-BUCKET2",
  "bucketCreatedAt": "2020-11-25T18:24:38+00:00",
  "bucketName": "DOC-EXAMPLE-BUCKET2",
  "allowsUnencryptedObjectUploads": "TRUE",
  "classifiableObjectCount": 8,
  "classifiableSizeInBytes": 133810,
  "jobDetails": {
    "isDefinedInJob": "TRUE",
    "isMonitoredByJob": "FALSE",
    "lastJobId": "188d4f6044d621771ef7d65f2example",
    "lastJobRunTime": "2021-04-09T19:37:11.511000+00:00"
  },
  "lastUpdated": "2021-04-30T07:33:06.337000+00:00",
  "objectCount": 8,
  "objectCountByEncryptionType": {
    "customerManaged": 0,
    "kmsManaged": 0,
    "s3Managed": 8,
    "unencrypted": 0,
    "unknown": 0
  },
  "publicAccess": {
    "effectivePermission": "NOT_PUBLIC",
    "permissionConfiguration": {
      "accountLevelPermissions": {
        "blockPublicAccess": {

```

```

        "blockPublicAcls": true,
        "blockPublicPolicy": true,
        "ignorePublicAcls": true,
        "restrictPublicBuckets": true
      }
    },
    "bucketLevelPermissions": {
      "accessControlList": {
        "allowsPublicReadAccess": false,
        "allowsPublicWriteAccess": false
      },
      "blockPublicAccess": {
        "blockPublicAcls": true,
        "blockPublicPolicy": true,
        "ignorePublicAcls": true,
        "restrictPublicBuckets": true
      },
      "bucketPolicy": {
        "allowsPublicReadAccess": false,
        "allowsPublicWriteAccess": false
      }
    }
  },
  "region": "us-east-1",
  "replicationDetails": {
    "replicated": false,
    "replicatedExternally": false,
    "replicationAccounts": []
  },
  "serverSideEncryption": {
    "kmsMasterKeyId": null,
    "type": "AES256"
  },
  "sharedAccess": "EXTERNAL",
  "sizeInBytes": 175978,
  "sizeInBytesCompressed": 0,
  "tags": [
    {
      "key": "Division",
      "value": "HR"
    },
    {
      "key": "Team",
      "value": "Recruiting"
    }
  ],
  "unclassifiableObjectCount": {
    "fileType": 0,
    "storageClass": 0,
    "total": 0
  },
  "unclassifiableObjectSizeInBytes": {
    "fileType": 0,
    "storageClass": 0,
    "total": 0
  },
  "versioning": true
}
]
}

```

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

```
{
```

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

### Example 1: Find buckets by bucket name

This example uses the [describe-buckets](#) command to query metadata for all buckets whose names begin with *my-S3* and are in the current AWS Region.

For Linux, macOS, or Unix:

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

For Microsoft Windows:

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

Where:

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

### Example 2: Find buckets that are publicly accessible

This example uses the [describe-buckets](#) command to query metadata for buckets that are in the current AWS Region and, based on a combination of permissions settings, are publicly accessible.

For Linux, macOS, or Unix:

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

For Microsoft Windows:

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

Where:

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

### Example 3: Find buckets that contain unencrypted objects

This example uses the [describe-buckets](#) command to query metadata for buckets that are in the current AWS Region and contain unencrypted objects.

For Linux, macOS, or Unix:

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

For Microsoft Windows:

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

Where:

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

## Example 4: Find buckets that aren't monitored by a job

This example uses the `describe-buckets` command to query metadata for buckets that are in the current AWS Region and aren't associated with any periodic sensitive data discovery jobs.

For Linux, macOS, or Unix:

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

For Microsoft Windows:

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

Where:

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

## Example 5: Find buckets that replicate data to external accounts

This example uses the `describe-buckets` command to query metadata for buckets that are in the current AWS Region and are configured to replicate objects to an AWS account that isn't part of your organization.

For Linux, macOS, or Unix:

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

For Microsoft Windows:

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

Where:

- `replicationDetails.replicatedExternally` specifies the JSON name of the **Replicated externally** field.

- `eq` specifies the *equals* operator.
- `true` specifies a Boolean value for the **Replicated externally** field.

### Example 6: Find buckets based on multiple criteria

This example uses the `describe-buckets` command to query metadata for buckets that are in the current AWS Region and meet the following criteria: are publicly accessible based on a combination of permission settings; contain unencrypted objects; and aren't associated with any periodic sensitive data discovery jobs.

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

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

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

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

Where:

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

## Allowing Amazon Macie to access S3 buckets and objects

When you enable Amazon Macie for your AWS account, Macie creates a [service-linked role \(p. 205\)](#) that grants Macie the permissions that it requires to call Amazon Simple Storage Service (Amazon S3) and other AWS services on your behalf. A service-linked role makes it easier to set up an AWS service because you don't have to manually add the necessary permissions for the service to complete actions on your behalf. To learn more about this type of role, see [Using service-linked roles](#) in the *AWS Identity and Access Management User Guide*.

The permissions policy for the Macie service-linked role allows Macie to perform actions that include retrieving information about your S3 buckets and objects, and retrieving objects from your S3 buckets. If

you're the Macie administrator for an organization, the policy also allows Macie to perform these actions for member accounts in your organization.

Macie uses these permissions to:

- Generate and maintain an inventory of your S3 buckets
- Provide statistical and other data about the buckets and objects in the buckets
- Monitor and evaluate the buckets for security and access control
- Analyze objects in the buckets to detect sensitive data

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

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

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

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

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

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

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

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

1. Sign in to the AWS Management Console and open the Amazon S3 console at <https://console.aws.amazon.com/s3/>.
2. In the navigation pane, choose **Buckets**.
3. Choose the bucket that you want to allow Macie to access.
4. On the **Permissions** tab, under **Bucket policy**, choose **Edit**.
5. In the **Bucket policy** editor, identify each `Deny` statement that restricts access and prevents Macie from accessing the bucket or the bucket's objects.

6. In each Deny statement, add a condition that uses the `aws:PrincipalArn` global condition key and specifies the ARN of the Macie service-linked role for your AWS account.

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

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

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

```
{
  "Version": "2012-10-17",
  "Id": "Policy1415115example",
  "Statement": [
    {
      "Sid": "Access to specific VPCE only",
      "Effect": "Deny",
      "Principal": "*",
      "Action": "s3:*",
      "Resource": [
        "arn:aws:s3::DOC-EXAMPLE-BUCKET",
        "arn:aws:s3::DOC-EXAMPLE-BUCKET/*"
      ],
      "Condition": {
        "StringNotEquals": {
          "aws:SourceVpce": "vpce-1a2b3c4d"
        }
      }
    }
  ]
}
```

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

```
{
  "Version": "2012-10-17",
  "Id": " Policy1415115example ",
  "Statement": [
    {
      "Sid": "Access-to-specific-VPCE-and-Macie-only",
      "Effect": "Deny",
      "Principal": "*",
      "Action": "s3:*",
      "Resource": [
        "arn:aws:s3::DOC-EXAMPLE-BUCKET",
        "arn:aws:s3::DOC-EXAMPLE-BUCKET/*"
      ],
      "Condition": {
        "StringNotEquals": {
          "aws:SourceVpce": "vpce-1a2b3c4d"
        },
        "StringNotLike": {

```

```
        "aws:PrincipalArn": "arn:aws:iam::123456789012:role/aws-service-role/  
macie.amazonaws.com/AWSServiceRoleForAmazonMacie"  
    }  
  }  
}  
]
```

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

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

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

For additional policy examples and detailed information about managing access to Amazon S3 resources, see [Identity and access management in Amazon S3](#) in the *Amazon Simple Storage Service User Guide*.



# Discovering sensitive data with Amazon Macie

To discover sensitive data with Amazon Macie, you create and run sensitive data discovery jobs. A sensitive data discovery job analyzes objects in Amazon Simple Storage Service (Amazon S3) buckets to determine whether the objects contain sensitive data, and it provides detailed reports of the sensitive data that it finds and the analysis that it performs. By creating and running sensitive data discovery jobs, you can automate discovery, logging, and reporting of sensitive data in your S3 buckets.

Each sensitive data discovery job can analyze objects by using managed data identifiers that Macie provides, custom data identifiers that you define, or a combination of the two. A *managed data identifier* is a set of built-in criteria and techniques that detect a specific type of sensitive data—for example, credit card numbers, AWS secret keys, or passport numbers for a particular country or region. These identifiers can detect a large and growing list of sensitive data types for many countries and regions, including multiple types of financial data, personal health information (PHI), and personally identifiable information (PII).

A *custom data identifier* is a set of criteria that you define to detect sensitive data. The criteria consist of a regular expression (*regex*) that defines a text pattern to match and, optionally, character sequences and a proximity rule that refine the results. By using custom data identifiers, you can detect sensitive data that reflects your organization's particular scenarios, intellectual property, or proprietary data—for example, employee IDs, customer account numbers, or internal data classifications.

When you create a job, you specify which S3 buckets contain objects that you want the job to analyze. Macie can analyze an object if the following is true:

- The object uses a supported file or storage format. For more information, see [Supported file and storage formats \(p. 89\)](#).
- If the object is encrypted, it's encrypted with a key that Macie is allowed to use. For more information, see [Analyzing encrypted S3 objects \(p. 90\)](#).
- The object is stored directly in Amazon S3 and uses a supported storage class—S3 Intelligent-Tiering, S3 One Zone-IA, S3 Standard, or S3 Standard-IA. Macie can't analyze data that's stored in Amazon S3 Glacier or other AWS services.

## Tip

Although Macie is optimized for Amazon S3, you can use it to discover sensitive data that you currently store elsewhere. You can do this by moving the data to Amazon S3 temporarily or permanently. For example, export Amazon RDS or Amazon Aurora snapshots to Amazon S3 in Apache Parquet format. Or export an Amazon DynamoDB table to Amazon S3. You can then create a job to analyze the data in Amazon S3.

- If the bucket has a restrictive bucket policy, the policy allows Macie to access objects in the bucket. For more information, see [Allowing Macie to access S3 buckets and objects \(p. 32\)](#).

As you create and configure a job, you also choose options to define the schedule and scope of the job's analysis. You can run a job only once, for on-demand analysis and assessment, or on a recurring basis for periodic analysis, assessment, and monitoring. To define the breadth and depth of a job's analysis, you can also choose various [scope options \(p. 57\)](#) for the job. These options include custom criteria that derive from properties of S3 buckets and objects, such as tags.

To help you meet and maintain compliance with your data security and privacy requirements, each sensitive data discovery job produces records of the sensitive data that it finds and the analysis that it performs—*sensitive data findings* and *sensitive data discovery results*. A *sensitive data finding* is a detailed

report of sensitive data that Macie found in an object. A *sensitive data discovery result* is a record that logs details about the analysis of an object. Each type of record adheres to a standardized schema, which can help you query, monitor, and process the records by using other applications, services, and systems as necessary. For more information, see [Reviewing job statistics and results \(p. 80\)](#).

#### Topics

- [Using managed data identifiers in Amazon Macie \(p. 37\)](#)
- [Building custom data identifiers in Amazon Macie \(p. 53\)](#)
- [Running sensitive data discovery jobs in Amazon Macie \(p. 56\)](#)
- [Supported file and storage formats in Amazon Macie \(p. 89\)](#)
- [Analyzing encrypted S3 objects with Amazon Macie \(p. 90\)](#)
- [Storing and retaining sensitive data discovery results with Amazon Macie \(p. 96\)](#)

## Using managed data identifiers in Amazon Macie

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

The topics in this section list and describe the categories and types of sensitive data that Macie can detect using managed data identifiers, and the relevant requirements for detecting that data.

#### Topics

- [Keyword requirements \(p. 37\)](#)
- [Sensitive data categories and types \(p. 38\)](#)
  - [Credentials \(p. 38\)](#)
  - [Financial information \(p. 39\)](#)
  - [Personal information – Personal health information \(p. 41\)](#)
  - [Personal information – Personally identifiable information \(p. 43\)](#)

For information about the types of data that Macie can analyze, see [Supported file and storage formats \(p. 89\)](#).

## Keyword requirements

To detect certain types of sensitive data, Macie requires a keyword to be in proximity of the data. If this is the case for a particular type of data, a subsequent topic in this section indicates specific keyword requirements for that data.

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

#### Structured, columnar data

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

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

#### Structured, record-based data

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

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

#### Unstructured data

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

Keywords aren't case sensitive. In addition, if a keyword contains a space, Macie automatically matches keyword variations that don't contain the space, or contain an underscore (`_`) or a hyphen (`-`) instead of the space.

## Sensitive data categories and types

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

- Credentials, for credentials data such as private keys or AWS secret keys.
- Financial information, for financial data such as credit card numbers or bank account numbers.
- Personal information, for personal health information (PHI) such as health insurance identification numbers, and personally identifiable information (PII) such as passport numbers.

Within each category, Macie can detect multiple types of sensitive data. The following topics list and describe each type and any relevant requirements for detecting it. For each type, they also indicate the unique identifier (**ID**) for the managed data identifier that's designed to detect the data. When you create a sensitive data discovery job, you can use this ID to explicitly include or exclude a managed data identifier from the job's analysis.

#### Topics

- [Credentials \(p. 38\)](#)
- [Financial information \(p. 39\)](#)
- [Personal information – Personal health information \(p. 41\)](#)
- [Personal information – Personally identifiable information \(p. 43\)](#)

## Credentials

The following table lists and describes the types of credentials that Macie can detect using managed data identifiers.

Detection type	ID	Keyword required	Additional information	Countries and regions
AWS secret key	AWS_CREDENTIALS	Yes, including: aws_secret_access_key, credentials, secret access key, secret key, set- awscredential	–	Any
OpenSSH private key	OPENSSSH_PRIVATE_KEY	Yes	–	Any
PGP private key	PGP_PRIVATE_KEY	No	–	Any
Public-Key Cryptography Standard (PKCS) private key	PKCS	No	–	Any
PuTTY private key	PUTTY_PRIVATE_KEY	No	–	Any

## Financial information

The following table lists and describes the types of financial information that Macie can detect using managed data identifiers. These are in addition to certain types of data that might also qualify as personally identifiable information (PII).

Detection type	ID	Keyword required	Additional information	Countries and regions
Bank account number	Depending on country or region: BANK_ACCOUNT_NUMBER (for Canadian and US bank account numbers), FRANCE_BANK_ACCOUNT_NUMBER, GERMANY_BANK_ACCOUNT_NUMBER, ITALY_BANK_ACCOUNT_NUMBER, SPAIN_BANK_ACCOUNT_NUMBER, UK_BANK_ACCOUNT_NUMBER	Yes, see <a href="#">the section called "Keywords for bank account numbers" (p. 40)</a>	This includes: <ul style="list-style-type: none"> <li>Canadian and US bank account numbers that consist of 6–19 digit sequences.</li> <li>International Bank Account Numbers (IBANs) that consist of up to 34 alphanumeric characters, including elements such as country code.</li> </ul>	Canada, France, Germany, Italy, Spain, UK, US
Credit card expiration date	CREDIT_CARD_EXPIRATION	Yes, including: expiration, expiry	Supported formats include MM/YY and YY/MM.	Any

Detection type	ID	Keyword required	Additional information	Countries and regions
Credit card magnetic strip data	CREDIT_CARD_MAGNETIC_STRIP	Yes, including: card data, iso7813, mag, magstripe, stripe, swipe	This includes tracks 1 and 2.	Any
Credit card number	CREDIT_CARD_NUMBER (for credit card numbers that are in proximity of a keyword) or CREDIT_CARD_NUMBER_(NO_KEYWORD) (for credit card numbers that aren't in proximity of a keyword)	Yes* May include:	Detection requires the data to be a 13–19 digit sequence that adheres to the Luhn check formula and uses a standard card number prefix for any of the following types of credit cards: American Express, Dankort, Diner's Club, Discover, Electron, Japanese Card Bureau (JCB), Mastercard, UnionPay, and Visa.	Any
Credit card verification code	CREDIT_CARD_SECURITY_CODE	Yes, including: card id, card identification code, card identification number, card security code, card validation code, card validation number, card verification data, card verification value, cvc, cvc2, cvv, cvv2, elo verification code	–	Any

\* Macie provides two managed data identifiers for credit card numbers, one for credit card numbers that are in proximity of a keyword (CREDIT\_CARD\_NUMBER) and another for credit card numbers that aren't in proximity of a keyword (CREDIT\_CARD\_NUMBER\_(NO\_KEYWORD)). For the former, required keywords include: account number, american express, amex, bank card, card num, card number, cc #, ccn, check card, credit card, credit card#, dankort, debit card, diners club, discover, electron, japanese card bureau, jcb, mastercard, mc, pan, payment account number, payment card number, pcn, union pay, visa.

## Keywords for bank account numbers

To detect various types of bank account numbers, Macie requires a keyword to be in proximity of the numbers. This includes Canadian and US bank account numbers that consist of 6–19 digit sequences.

This also includes International Bank Account Numbers (IBANs) that consist of up to 34 alphanumeric characters, including elements such as country code.

The following table lists the keywords that Macie recognizes for specific countries and regions.

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

## Personal information – Personal health information

The following table lists and describes the types of personal health information (PHI) that Macie can detect using managed data identifiers. These are in addition to certain types of data that might also qualify as personally identifiable information (PII).

Detection type	ID	Keyword required	Additional information	Countries and regions
Drug Enforcement Agency (DEA) Registration Number	US_DRUG_ENFORCEMENT_AGENCY_NUMBER	DEA, dea, dea number, dea registration		US

Detection type	ID	Keyword required	Additional information	Countries and regions
Health Insurance Claim Number (HICN)	USA_HEALTH_INSURANCE_CLAIM_NUMBER	Yes, including: health insurance claim number, hic no, hic no., hic number, hic#, hcn, hcn#., hcnno#		US
Health insurance or medical identification number	Depending on country or region: CANADA_HEALTH_INSURANCE_CARD_NUMBER, EUROPEAN_HEALTH_INSURANCE_CARD_NUMBER, FINLAND_EUROPEAN_HEALTH_INSURANCE_CARD_NUMBER, FRANCE_HEALTH_INSURANCE_NUMBER, UK_NHS_NUMBER, USA_MEDICARE_BENEFICIARY_IDENTIFIER	Yes, see <a href="#">the section called "Keywords for health insurance numbers" (p. 42)</a>	This includes European Health Insurance Card numbers (EU, Finland), health insurance numbers (France), Medicare Beneficiary Identifiers (US), NHS numbers (UK), and Personal Health Numbers (Canada).	Canada, EU, Finland, France, UK, US
Healthcare Common Procedure Coding System (HCPCS) code	USA_HEALTHCARE_PROCEDURE_CODE	Yes, including: current procedural terminology, hcpcs, healthcare common procedure coding system	–	US
National Drug Code (NDC)	USA_NATIONAL_DRUG_CODE	Yes, including: national drug code, ndc	–	US
National Provider Identifier (NPI)	USA_NATIONAL_PROVIDER_IDENTIFIER	Yes, including: hipaa, n.p.i, national provider, npi	–	US
Unique device identifier (UDI)	MEDICAL_DEVICE_UID	Yes, including: blood, blood bag, dev id, device id, device identifier, gs1, hibcc, iccbba, med, udi, unique device id, unique device identifier	As defined by the US Food and Drug Administration's Global Unique Device Identification Database (GUDID).	US

## Keywords for health insurance and medical identification numbers

To detect various types of health insurance and medical identification numbers, Macie requires a keyword to be in proximity of the numbers. This includes European Health Insurance Card numbers (EU, Finland), health insurance numbers (France), Medicare Beneficiary Identifiers (US), National Insurance numbers (UK), NHS numbers (UK), and Personal Health Numbers (Canada).

The following table lists the keywords that Macie recognizes for specific countries and regions.

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

## Personal information – Personally identifiable information

The following table lists and describes the types of personally identifiable information (PII) that Macie can detect using managed data identifiers.

Detection type	ID	Keyword required	Additional information	Countries and regions
Birth date	DATE_OF_BIRTH	Yes, including: bday, b-day, birth date, birthday, date of birth, dob	This includes most date formats, such as all digits, and combinations of digits and names	Any



Detection type	ID	Keyword required	Additional information	Countries and regions
			of months. Date components can be separated by spaces, slashes (/), or hyphens (-).	
Driver's license identification number	Depending on country or region: AUSTRALIA_DRIVERS_LICENSE, AUSTRIA_DRIVERS_LICENSE, BELGIUM_DRIVERS_LICENSE, BULGARIA_DRIVERS_LICENSE, CANADA_DRIVERS_LICENSE, CROATIA_DRIVERS_LICENSE, CYPRUS_DRIVERS_LICENSE, CZECHIA_DRIVERS_LICENSE, DENMARK_DRIVERS_LICENSE, DRIVERS_LICENSE (for the US), ESTONIA_DRIVERS_LICENSE, FINLAND_DRIVERS_LICENSE, FRANCE_DRIVERS_LICENSE, GERMANY_DRIVERS_LICENSE, GREECE_DRIVERS_LICENSE, HUNGARY_DRIVERS_LICENSE, IRELAND_DRIVERS_LICENSE, ITALY_DRIVERS_LICENSE, LATVIA_DRIVERS_LICENSE, LITHUANIA_DRIVERS_LICENSE, LUXEMBOURG_DRIVERS_LICENSE, MALTA_DRIVERS_LICENSE, NETHERLANDS_DRIVERS_LICENSE, POLAND_DRIVERS_LICENSE, PORTUGAL_DRIVERS_LICENSE, ROMANIA_DRIVERS_LICENSE, SLOVAKIA_DRIVERS_LICENSE, SLOVENIA_DRIVERS_LICENSE, SPAIN_DRIVERS_LICENSE, SWEDEN_DRIVERS_LICENSE, UK_DRIVERS_LICENSE	Yes, see <a href="#">the section called "Licenses for license identification" (p. 49)</a>	–	Australia, Austria, Belgium, Bulgaria, Canada, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, UK, US
Electoral roll number	UK_ELECTORAL_ROLL_NUMBER	Yes, including: electoral #, electoral number, electoral roll #, electoral roll no., electoral roll number, electoralrollno	–	UK

Detection type	ID	Keyword required	Additional information	Countries and regions
Full name	NAME	No	Macie can detect full names only. Support is limited to Latin character sets.	Any
Global Positioning System (GPS) coordinates	LATITUDE_LONGITUDE	Yes, including: coordinate, coordinates, lat long, latitude longitude, position	<p>Macie can detect GPS coordinates only if the latitude and longitude coordinates are stored as a pair and they're in Decimal Degrees (DD) format, for example: 41.948614, -87.655311.</p> <p>Support doesn't include coordinates in Degrees Decimal Minutes (DDM) format, for example 41° 56.9168 'N 87° 39.3187 'W, or Degrees, Minutes, Seconds (DMS) format, for example 41° 56 ' 55.0104 "N 87° 39 ' 19.1196 "W.</p>	Australia, Canada, Ireland, UK, US
Mailing address	ADDRESS or BRAZIL_CEP_CODE (for Brazil's Código de Endereçamento Postal)	No	Although a keyword isn't required, detection requires the address to include the name of a city or place.	Australia, Canada, France, Germany, Italy, Spain, UK, US



Detection type	ID	Keyword required	Additional information	Countries and regions
Permanent residence number	CANADA_NATIONAL_IDENTIFICATION_NUMBER	Yes, including: carte résident permanent, numéro carte résident permanent, numéro résident permanent, permanent resident card, permanent resident card number, permanent resident no., permanent resident no., permanent resident number, pr no, pr no., pr non, pr number, résident permanent no., résident permanent non		Canada
Phone number	Depending on country or region: BRAZIL_PHONE_NUMBER, FRANCE_PHONE_NUMBER, GERMANY_PHONE_NUMBER, ITALY_PHONE_NUMBER, PHONE_NUMBER (for Canada and the US), SPAIN_PHONE_NUMBER, UK_PHONE_NUMBER	Yes, including: cel, cell, celular, cell, mobile, número residencial, numero residencial, phone, phone number, telephone, telephone number	This includes toll-free numbers in the US and fax numbers.  If a keyword is in proximity of the data, the number doesn't have to include a country code.  If a keyword isn't in proximity of the data, the number has to include a country code.	Brazil, Canada, France, Germany, Italy, Spain, UK, US

Detection type	ID	Keyword required	Additional information	Countries and regions
Social Insurance Number (SIN)	CANADA_SOCIAL_INSURANCE_NUMBER	Yes, including: canadian id, numéro d'assurance sociale, sin, social insurance number  Also refer to <a href="#">the section called "Keywords for health insurance and medical identification numbers"</a> (p. 42)	–	Canada
Social Security number (SSN)	Depending on country or region: SPAIN_SOCIAL_SECURITY_NUMBER or USA_SOCIAL_SECURITY_NUMBER	Yes, including: • Spain – número de la seguridad social, social security no., social security number, social security no., ssn, ssn# • US – social security number, ss#, ssn	–	Spain, US
Taxpayer identification or reference number	Depending on country or region: AUSTRALIA_TAX_FILE_NUMBER, BRAZIL_CNPJ_NUMBER, BRAZIL_CPF_NUMBER, FRANCE_TAX_IDENTIFICATION_NUMBER, GERMANY_TAX_IDENTIFICATION_NUMBER, SPAIN_NIE_NUMBER, SPAIN_NIF_NUMBER, SPAIN_TAX_IDENTIFICATION_NUMBER, UK_TAX_IDENTIFICATION_NUMBER, USA_INDIVIDUAL_TAX_IDENTIFICATION_NUMBER	Yes, see <a href="#">the section called "Keywords for taxpayer identification numbers"</a> (p. 42)	This includes: CIF, NIE, and NIF (Spain); CNPJ and CPF (Brazil); Codice Fiscale (Italy); ITIN (US); Steueridentifikationsnummer (Germany); TFN (Australia); TIN (France); and, TRN, UTR (UK).	Australia, Brazil, France, Germany, Italy, Spain, UK, US

Detection type	ID	Keyword required	Additional information	Countries and regions
Vehicle identification number (VIN)	VEHICLE_IDENTIFICATION_NUMBER	Yes, <b>VEHICLE_IDENTIFICATION_NUMBER</b> Fahrgestellnummer, nív, numărul de identificare, numărul seriei de sasiu, numer VIN, Número de Identificação do Veículo, Número de Identificación de Automóviles, numéro d'identification du véhicule, vehicle identification number, vin, VIN numeris	Macie can detect VINs that consist of a 17-character sequence and adhere to the ISO 3779 and 3780 standards. These standards were designed for worldwide use.	Any, if the VIN is in proximity of a keyword in one of the following languages: English, French, German, Lithuanian, Polish, Portuguese, Romanian, or Spanish

## Keywords for driver's license identification numbers

To detect various types of driver's license identification numbers, Macie requires a keyword to be in proximity of the numbers. The following table lists the keywords that Macie recognizes for specific countries and regions.

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

Country or region	Keywords
	licence, driving license, driving permit, permis de conduire
Croatia	vozačka dozvola
Cyprus	άδεια οδήγησης
Czech Republic	číslo licence, číslo licence řidiče, číslo řidičského průkazu, ovladače lic., povolení k jízdě, povolení řidiče, řidiči povolení, řidičský průkaz, řidičský průkaz
Denmark	kørekort, kørekortnummer
Estonia	juhi litsentsi number, juhiloa number, juhiluba, juhiluba number
Finland	ajokortin numero, ajokortti, förare lic., körkort, körkort nummer, kuljettaja lic., permis de conduire
France	permis de conduire
Germany	fuehrerschein, fuehrerschein- nr, fuehrerscheinnummer, fuhrerschein, fuhrerschein, fuhrerschein- nr, fuhrerschein- nr, fuhrerscheinnummer, fuhrerscheinnummer
Greece	δεια οδήγησης, adeia odigisis
Hungary	illesztőprogramok lic, jogosítvány, jogsí, licencszám, vezető engedély, vezetői engedély
Ireland	ceadúnas tiomána
Italy	patente di guida, patente di guida numero, patente guida, patente guida numero
Latvia	autovadītāja apliecība, licences numurs, vadītāja apliecība, vadītāja apliecības numurs, vadītāja atļauja, vadītāja licences numurs, vadītāji lic.
Lithuania	vairuotojo pažymėjimas
Luxembourg	fahrerlaubnis, fuhrerschäin
Malta	liċenzja tas-sewqan
Netherlands	permis de conduire, rijbewijs, rijbewijsnummer
Poland	numer licencyjny, prawo jazdy, zezwolenie na prowadzenie
Portugal	carta de condução, carteira de habilitação, carteira de motorist, carteira habilitação, carteira motorist, licença condução, licença de condução, número de licença, número licença, permissão condução, permissão de condução

Country or region	Keywords
Romania	numărul permisului de conducere, permis de conducere
Slovakia	číslo licencie, číslo vodičského preukazu, ovládače lic., povolenia vodičov, povolenie jazdu, povolenie na jazdu, povolenie vodiča, vodičský preukaz
Slovenia	vozniško dovoljenje
Spain	carnet conductor, el carnet de conducir, licencia conducir, licencia de manejo, número carnet conducir, número de carnet de conducir, número de permiso conducir, número de permiso de conducir, número licencia conducir, número permiso conducir, permiso conducción, permiso conducir, permiso de conducción
Sweden	ajokortin numero, dlno# ajokortti, drivere lic., förare lic., körkort, körkort nummer, körkortsnummer, kuljettajat lic. דרייווערס דערלויבעניש , שאפער דערלויבעניש נומער
UK	dl#, dl:, dlno#, driver licence, driver licences, driver license, driver licenses, driver permit, drivers lic., drivers licence, driver's licence, drivers licences, driver's licences, drivers license, driver's license, drivers licenses, driver's licenses, drivers permit, driver's permit, drivers permit number, driving licence, driving license, driving permit
US	dl#, dl:, dlno#, driver licence, driver licences, driver license, driver licenses, driver permit, drivers lic., drivers licence, driver's licence, drivers licences, driver's licences, drivers license, driver's license, drivers licenses, driver's licenses, drivers permit, driver's permit, drivers permit number, driving licence, driving license, driving permit

## Keywords for national identification numbers

To detect various types of national identification numbers, Macie requires a keyword to be in proximity of the numbers. This includes Documento Nacional de Identidad (DNI) identifiers (Spain), French National Institute for Statistics and Economic Studies (INSEE) codes, German National Identity Card numbers, and Registro Geral (RG) numbers (Brazil).

The following table lists the keywords that Macie recognizes for specific countries and regions.

Country or region	Keywords
Brazil	registro geral, rg
France	assurance sociale, carte nationale d'identité, cni, code sécurité sociale, French social security number, fssn#, insee, insurance number, national



Country or region	Keywords
	id number, nationalid#, numéro d'assurance, sécurité sociale, sécurité sociale non., sécurité sociale numéro, social, social security, social security number, socialsecuritynumber, ss#, ssn, ssn#
Germany	ausweisnummer, id number, identification number, identity number, insurance number, personal id, personalausweis
Italy	codice fiscale, dati anagrafici, ehic, health card, health insurance card, p. iva, partita i.v.a., personal data, tax code, tessera sanitaria
Spain	dni, dni#, dninúmero#, documento nacional de identidad, identidad único, identidadúnico#, insurance number, national identification number, national identity, nationalid#, nationalidno#, número nacional identidad, personal identification number, personal identity no, unique identity number, uniqueid#

## Keywords for passport numbers

To detect various types of passport numbers, Macie requires a keyword to be in proximity of the numbers. The following table lists the keywords that Macie recognizes for specific countries and regions.

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

## Keywords for taxpayer identification and reference numbers

To detect various types of taxpayer identification and reference numbers, Macie requires a keyword to be in proximity of the numbers. The following table lists the keywords that Macie recognizes for specific countries and regions.

Country or region	Keywords
Australia	tax file number, tfn
Brazil	cadastro de pessoa física, cadastro de pessoa física, cadastro de pessoas físicas, cadastro de pessoas físicas, cadastro nacional da pessoa jurídica, cadastro nacional da pessoa jurídica, cnpj, cpf
France	numéro d'identification fiscal, tax id, tax identification number, tax number, tin, tin#
Germany	identifikationsnummer, steuer id, steueridentifikationsnummer, steuernummer, tax id, tax identification number, tax number
Italy	codice fiscale, dati anagrafici, ehic, health card, health insurance card, p. iva, partita i.v.a., personal data, tax code, tessera sanitaria
Spain	cif, cif número, cifnúmero#, nie, nif, número de contribuyente, número de identidad de extranjero, número de identificación fiscal, número de impuesto corporativo, personal tax number, tax id, tax identification number, tax number, tin, tin#
UK	paye, tax id, tax id no., tax id number, tax identification, tax identification#, tax no., tax number, tax reference, tax#, taxid#, temporary reference number, tin, trn, unique tax reference, unique taxpayer reference, utr
US	i.t.i.n., individual taxpayer identification number, itin

## Building custom data identifiers in Amazon Macie

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

With custom data identifiers, you can define detection rules that reflect your organization's particular scenarios, intellectual property, or proprietary data—for example, employee IDs, customer account numbers, or internal data classifications. By using these identifiers in sensitive data discovery jobs, you can perform targeted analysis of your organization's Amazon Simple Storage Service (Amazon S3) data in a way that supplements the [managed data identifiers \(p. 37\)](#) that Amazon Macie provides.

### Topics

- [Components of a custom data identifier \(p. 54\)](#)

- [Creating custom data identifiers \(p. 55\)](#)
- [Regex support in custom data identifiers \(p. 56\)](#)

## Components of a custom data identifier

When you create a custom data identifier, you specify a regular expression (*regex*) that defines a text pattern to match in data. The regex can contain as many as 500 characters.

You can also specify certain character sequences, such as words and phrases, and a proximity rule to refine your analysis of data.

### Keywords

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

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

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

### Maximum match distance

This is the maximum number of characters that can exist between text that matches the regex pattern and one or more of the keywords that you specify. If text matches the regex pattern and is within the specified distance from a keyword, Macie reports that occurrence of the text.

You can specify a distance of 1–300 characters. The default distance is 50 characters.

### Ignore words

These are specific character sequences to exclude from the results. If text matches the regex pattern and it contains an ignore word, Macie doesn't report that occurrence of the text.

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

For example, many companies have a specific syntax for employee IDs. One such syntax might be: a capital letter that indicates whether the employee is a full-time (F) or part-time (P) employee, followed by a hyphen (-), followed by an eight-digit sequence that identifies the employee. Examples are: F-12345678, for a full-time employee, and P-87654321, for a part-time employee.

If you create a custom data identifier to detect employee IDs that use this syntax, you might use the following regex: `[A-Z]-\d{8}`. To refine the analysis and avoid false positives, you might also configure

the custom data identifier to report only those instances where the keyword `employee` is within a specific distance of text that matches the regex pattern.

## Creating custom data identifiers

The following steps explain how to create a custom data identifier by using the Amazon Macie console.

### To create a custom data identifier

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

We strongly recommend that you avoid including any sensitive data in this name. Other users of your account might be able to see the identifier's name, depending on the actions that they're allowed to perform in Macie.

5. For **Description**, enter a brief description of the custom data identifier. The description can contain as many as 512 characters.

We strongly recommend that you avoid including any sensitive data in the description. Other users of your account might be able to see the identifier's description, depending on the actions that they're allowed to perform in Macie.

6. For **Regular expression**, enter the regular expression (*regex*) that defines the pattern to match. The regex can contain as many as 500 characters. To learn about supported syntax and constraints, see [Regex support in custom data identifiers \(p. 56\)](#) later in this section.
7. (Optional) For **Keywords**, enter as many as 50 character sequences (separated by commas) that define specific text to match. Each keyword can contain 3–90 UTF-8 characters. Keywords aren't case sensitive.

Macie includes a result if the text matches the regex pattern and is within the maximum match distance of one of these keywords, as explained in the [preceding topic \(p. 54\)](#).

8. (Optional) For **Ignore words**, enter up to 10 character sequences (separated by commas) that define specific text to exclude from the results. Each ignore word can contain 4–90 UTF-8 characters. Ignore words are case sensitive.

Macie excludes results for text that contains any of these words, even if the text matches the regex pattern.

9. (Optional) For **Maximum match distance**, enter the maximum allowable distance between text that matches the regex pattern and any of the keywords. The default distance is 50 characters.

Macie includes a result only if the text matches the regex pattern and is within this distance of a keyword, as explained in the [preceding topic \(p. 54\)](#).

10. (Optional) Test the custom data identifier by pasting up to 1,000 characters of text into the **Sample data** box, and then choosing **Submit**. Macie evaluates the sample data by using the identifier, and reports the number of matches. You can repeat this step as many times as you like to refine and optimize the identifier.

#### Note

We highly recommend that you test and refine the custom data identifier before you save it. Because custom data identifiers are used by sensitive data discovery jobs, you can't edit a custom data identifier after you save it. This helps ensure that you have an immutable history of sensitive data findings and discovery results for data privacy and protection audits or investigations that you perform.

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

After you create a custom data identifier, you can use it to analyze objects in Amazon Simple Storage Service (Amazon S3) buckets by [creating and running a sensitive data discovery job \(p. 65\)](#). When you create a job, you optionally specify one or more custom data identifiers that you want the job to use when it analyzes data.

## Regex support in custom data identifiers

Macie supports a subset of the regex pattern syntax provided by the [Perl Compatible Regular Expressions \(PCRE\) library](#).

Of the constructs provided by the PCRE library, Macie doesn't support the following pattern elements:

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

To protect against malformed or long-running expressions, Macie automatically tests custom data identifiers against a collection of sample text.

The following tips and recommendations can help you create effective regex patterns for custom data identifiers in Macie:

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

## Running sensitive data discovery jobs in Amazon Macie

With Amazon Macie, you create and run sensitive data discovery jobs to automate discovery, logging, and reporting of sensitive data in Amazon Simple Storage Service (Amazon S3) buckets. A sensitive data discovery job analyzes objects in S3 buckets to determine whether the objects contain sensitive data, and it provides detailed reports of the sensitive data that it finds and the analysis that it performs.

To help you meet and maintain compliance with your data security and privacy requirements, Macie provides several options for scheduling and defining the scope of each job. With these options, you can

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

You can configure a job to run only once for on-demand analysis and assessment, or on a recurring basis for periodic analysis, assessment, and monitoring. In addition, you define the breadth and depth of each job's analysis. When you create a job, you start by specifying which S3 buckets you want the job to analyze—specific buckets that you select or buckets that match specific criteria. You can then refine the scope of that analysis by choosing various options, including custom include and exclude criteria that derive from properties of S3 objects. You can also specify the types of sensitive data that you want the job to detect. A job can analyze objects by using the [managed data identifiers \(p. 37\)](#) that Macie provides, [custom data identifiers \(p. 53\)](#) that you define, or a combination of the two. By selecting specific managed and custom data identifiers for a job, you can tailor the job's analysis to focus on specific types of sensitive data.

Each job produces records of the sensitive data that it finds and the analysis that it performs—*sensitive data findings* and *sensitive data discovery results*. A *sensitive data finding* is a detailed report of sensitive data that Macie found in an object. A *sensitive data discovery result* is a record that logs details about the analysis of an object. Macie creates a sensitive data discovery result for each object that you configure a job to analyze, including objects that don't contain sensitive data. Each type of record adheres to a standardized schema, which can help you query, monitor, and process the records to meet your security and compliance requirements.

#### Topics

- [Scope options for sensitive data discovery jobs \(p. 57\)](#)
- [Creating a sensitive data discovery job \(p. 65\)](#)
- [Monitoring sensitive data discovery jobs with Amazon CloudWatch Logs \(p. 71\)](#)
- [Reviewing statistics and results for a sensitive data discovery job \(p. 80\)](#)
- [Managing sensitive data discovery jobs \(p. 82\)](#)
- [Forecasting and monitoring costs for sensitive data discovery jobs \(p. 87\)](#)

## Scope options for sensitive data discovery jobs

In Amazon Macie, you define the scope of the data that a sensitive data discovery job analyzes. To help you do this, Macie provides several job-specific options that you can choose when you create and configure a job.

#### Scope options

- [S3 buckets \(p. 57\)](#)
- [Include existing S3 objects \(p. 62\)](#)
- [Sampling depth \(p. 62\)](#)
- [S3 object criteria \(p. 63\)](#)

## S3 buckets

The first step in creating a sensitive data discovery job is to specify which Amazon Simple Storage Service (Amazon S3) buckets contain objects that you want the job to analyze. You can do this in either of two ways, by selecting specific S3 buckets from your bucket inventory or by specifying custom criteria that derive from properties of S3 buckets.

#### Selecting specific buckets

With this option, you explicitly select each S3 bucket that you want the job to analyze. Then, when the job runs, it analyzes objects in the selected buckets. If you also configure the job to run

periodically on a daily, weekly, or monthly basis, the job analyzes objects in those same buckets each time it runs.

This configuration is helpful for cases where you prefer to perform targeted analysis of a specific set of data. It gives you precise, predictable control over which buckets a job analyzes.

### Specifying bucket criteria

With this option, you define runtime criteria that determine which S3 buckets the job analyzes. The criteria consist of one or more conditions that derive from bucket properties, such as public access settings and tags. When the job runs, it identifies buckets that match your criteria and then analyzes objects in those buckets. If you also configure the job to run periodically, the job does this each time it runs. Consequently, the job might analyze objects in different buckets each time it runs, depending on changes to your bucket inventory and the criteria that you define.

This configuration is helpful for cases where you want the scope of the job's analysis to dynamically adapt to changes to your bucket inventory. For example, if you configure a job to use bucket criteria and run periodically, the job can automatically identify new buckets that match the criteria and inspect those buckets for sensitive data.

The topics in this section provide additional details about each option.

### Topics

- [Selecting S3 buckets \(p. 58\)](#)
- [Specifying S3 bucket criteria \(p. 60\)](#)

## Selecting S3 buckets

If you choose to explicitly select each S3 bucket that you want a job to analyze, Macie provides you with a complete inventory of your buckets in the current AWS Region. You can then review your inventory and select the buckets that you want. To learn how Macie generates and maintains this inventory for you, see [How Macie monitors Amazon S3 data \(p. 8\)](#).

If your account is the Macie administrator account for an organization, the inventory includes buckets that are owned by member accounts in your organization. You can select as many as 1,000 of these buckets, spanning as many as 1,000 accounts.

To help you make your bucket selections, the inventory provides details and statistics for each bucket. This includes the amount of data that a job can analyze in each bucket—*classifiable* objects are objects that use a supported Amazon S3 storage class (S3 Intelligent-Tiering, S3 One Zone-IA, S3 Standard, or S3 Standard-IA) and have a file name extension for a [supported file or storage format \(p. 89\)](#). The inventory also indicates whether any existing jobs are configured to analyze objects in a bucket. These details can help you estimate the breadth of a job and refine your bucket selections.

In the inventory table:



- **Classifiable objects** – This field indicates the total number of objects that the job can analyze in a bucket.
- **Classifiable size** – This field indicates the total storage size of all the objects that the job can analyze in a bucket.


If a bucket contains compressed objects, this value doesn't reflect the actual size of those objects after they're decompressed. If versioning is enabled for a bucket, this value is based on the storage size of the latest version of each object in the bucket.

- **Monitored** – This field indicates whether any existing jobs are configured to periodically analyze objects in a bucket on a daily, weekly, or monthly basis.

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

- **Latest job run** – If any existing periodic or one-time jobs are configured to analyze objects in a bucket, this field indicates the most recent time when one of those jobs started to run. Otherwise, this field is empty.

If the information icon () appears next to any bucket names in the table, we recommend that you retrieve the latest bucket metadata from Amazon S3. To do this, choose refresh () above the table. The information icon indicates that a bucket was created during the past 24 hours, possibly after Macie last retrieved bucket and object metadata from Amazon S3 as part of the daily refresh cycle. For more information, see [Data refreshes \(p. 10\)](#).




If the warning icon () appears next to a bucket's name in the table, Macie isn't allowed to access the bucket or the bucket's objects. (Macie can only provide a subset of information about the bucket, such as the bucket's name.) This means that the job won't be able to analyze objects in the bucket. To investigate the issue, review the bucket's policy and permissions settings in Amazon S3. For example, the bucket might have a restrictive bucket policy. For more information, see [Allowing Macie to access S3 buckets and objects \(p. 32\)](#).

To customize your view of the inventory and find specific buckets more easily, you can filter the table by entering filter criteria in the filter bar. The following table provides some examples.

To show all buckets that...	Apply this filter...
Are owned by a specific account	<b>Account ID</b> = <i>the 12-digit ID for the account</i>
Are publicly accessible	<b>Effective permission</b> = <b>PUBLIC</b>
Aren't included in any periodic jobs	<b>Actively monitored by job</b> = <b>FALSE</b>
Aren't included in any periodic or one-time jobs	<b>Defined in job</b> = <b>FALSE</b>
Have a specific tag key*	<b>Tag key</b> = <i>the tag key</i>
Have a specific tag value*	<b>Tag value</b> = <i>the tag value</i>
Contain unencrypted objects (or use client-side encryption)	<b>Object count by encryption</b> is <b>No encryption</b> and <b>From</b> = <b>1</b>

\* Tag keys and values are case sensitive. Also, you have to specify a complete, valid value for these fields in a filter. You can't specify partial values or use wildcard characters.

To display the details of a bucket, choose the bucket's name and refer to the details panel. From there, you can also:

- Pivot and drill down on certain fields by choosing a magnifying glass for the field. Choose  to show buckets with the same value, or choose  to show buckets with other values.
- Retrieve the latest metadata for objects in the bucket. This can be helpful if you recently created a bucket or made significant changes to the bucket's objects during the past 24 hours. To retrieve the data, choose refresh () in the **Object statistics** section of the panel. This option is available for buckets that contain 30,000 or fewer objects.



## Specifying S3 bucket criteria

If you choose to specify bucket criteria for a job, Macie provides options for defining and testing the criteria. These are runtime criteria that determine which S3 buckets contain objects for the job to analyze. Each time the job runs, it identifies buckets that match your criteria and then analyzes objects in the appropriate buckets.

### Defining bucket criteria

Bucket criteria consist of one or more conditions that derive from properties of S3 buckets. Each condition, also referred to as a *criterion*, consists of the following parts:

- A property-based field, such as **Account ID** or **Effective permission**.
- An operator, either *equals* (eq) or *not equals* (neg).
- One or more values.
- An include or exclude statement that indicates whether you want the job to analyze (*include*) or skip (*exclude*) buckets that match the condition.

If you specify more than one value for a field, Macie uses OR logic to join the values. If you specify more than one condition for the criteria, Macie uses AND logic to join the conditions. In addition, exclude conditions take precedence over include conditions. For example, if you include buckets that are publicly accessible and exclude buckets that have specific tags, the job analyzes objects in any bucket that's publicly accessible unless the bucket has one of the specified tags.

You can define conditions that derive from any of the following property-based fields for S3 buckets.

#### Account ID

The unique identifier for the AWS account that owns a bucket. To specify multiple values for this field, enter the ID for each account and separate each entry with a comma.

Note that values are case sensitive. In addition, Macie doesn't support use of wildcard characters or partial values for this field.

#### Bucket name

The name of a bucket. This field correlates to the **Name** field, not the **Amazon Resource Name (ARN)** field, in Amazon S3. To specify multiple values for this field, enter the name of each bucket and separate each entry with a comma.

Note that values are case sensitive. In addition, Macie doesn't support use of wildcard characters or partial values for this field.

#### Effective permission

Specifies whether a bucket is publicly accessible. You can choose one or more of the following values for this field:

- **NOT\_PUBLIC** – The general public doesn't have read or write access to the bucket.
- **PUBLIC** – The general public has read or write access to the bucket.
- **UNKNOWN** – Macie wasn't able to evaluate the public access settings for the bucket.

To determine this value for a bucket, Macie analyzes a combination of account- and bucket-level settings for the bucket: the block public access setting for the account; the block public access setting for the bucket; the bucket policy for the bucket; and, the access control list (ACL) for the bucket.

#### Shared access

Specifies whether a bucket is shared with other AWS accounts. You can choose one or more of the following values for this field:

- **EXTERNAL** – The bucket is shared with accounts that aren't in the same organization.
- **INTERNAL** – The bucket is shared with accounts in the same organization.
- **NOT\_SHARED** – The bucket isn't shared with other accounts.
- **UNKNOWN** – Macie wasn't able to evaluate the shared access settings for the bucket.

To determine this value for a bucket, Macie analyzes the bucket policy and ACL for the bucket. In addition, an *organization* is defined as a set of Macie accounts that are centrally managed as a group of related accounts through AWS Organizations or by Macie invitation.

## Tags

The tags that are associated with a bucket. Tags are custom labels that you can associate with specific types of AWS resources, including S3 buckets. Each tag consists of a required tag key and an optional tag value. For information about tagging S3 buckets, see [Categorizing your storage using tags](#) in the *Amazon Simple Storage Service User Guide*.

For a sensitive data discovery job, you can use this type of condition to include or exclude buckets that have a specific tag key, a specific tag value, or a specific tag key and tag value (as a pair). For example:

- If you specify **Project** as a tag key and don't specify any tag values for a condition, any bucket that has the *Project* tag key meets the condition's criteria, regardless of the tag values that are associated with that tag key.
- If you specify **Development** and **Test** as tag values and don't specify any tag keys for a condition, any bucket that has the **Development** or **Test** tag value meets the condition's criteria, regardless of the tag keys that are associated with those tag values.

To specify multiple tag keys in a condition, enter each tag key in the **Key** field and separate each entry with a comma. To specify multiple tag values in a condition, enter each tag value in the **Value** field and separate each entry with a comma.

Note that tag keys and values are case sensitive. In addition, Macie doesn't support use of wildcard characters or partial values in tag conditions.

## Testing bucket criteria

While you define your bucket criteria, you can test and refine the criteria by previewing the results. To do this, expand the **Preview the criteria results** section that appears below the criteria on the console. This section displays a table of all the buckets that currently match the criteria.

The table also provides insight into the amount of data that the job can analyze in each bucket—*classifiable* objects are objects that use a supported Amazon S3 storage class (S3 Intelligent-Tiering, S3 One Zone-IA, S3 Standard, or S3 Standard-IA) and have a file name extension for a [supported file or storage format](#) (p. 89). The table also indicates whether any existing jobs are configured to periodically analyze objects in a bucket.

In the table:

- **Classifiable objects** – This field indicates the total number of objects that the job can analyze in a bucket.
- **Classifiable size** – This field indicates the total storage size of all the objects that the job can analyze in a bucket.

If a bucket contains compressed objects, this value doesn't reflect the actual size of those objects after they're decompressed. If versioning is enabled for a bucket, this value is based on the storage size of the latest version of each object in the bucket.

- **Monitored** – This field indicates whether any existing jobs are configured to periodically analyze objects in a bucket on a daily, weekly, or monthly basis.

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

If the warning icon (⚠) appears next to a bucket's name, Macie isn't allowed to access the bucket or the bucket's objects. (Macie can only provide a subset of information about the bucket, such as the bucket's name.) This means that the job won't be able to analyze objects in the bucket. To investigate the issue, review the bucket's policy and permissions settings in Amazon S3. For example, the bucket might have a restrictive bucket policy. For more information, see [Allowing Macie to access S3 buckets and objects](#) (p. 32).

To refine the bucket criteria for the job, use the filter settings to add, change, or remove conditions from the criteria. Macie then updates the table to reflect your changes.

## Include existing S3 objects

You can use sensitive data discovery jobs to perform ongoing, incremental analysis of objects in S3 buckets. If you configure a job to run periodically, Macie does this for you automatically—each run analyzes only those objects that are created or changed after the preceding run. With the **Include existing objects** option, you choose the starting point for the first increment:

- To analyze all existing objects immediately after you finish creating the job, select the check box for this option.
- To wait and analyze only those objects that are created or changed after you create the job and before the first run, clear the check box for this option.

Clearing this check box is helpful for cases where you've already analyzed the data and want to continue to analyze it periodically. For example, if you previously used Amazon Macie Classic to classify data and you recently moved to Macie, you might use this option to ensure continued discovery and classification of your data without incurring unnecessary costs or duplicating classification data.

Each subsequent run of a periodic job automatically analyzes only those objects that are created or changed after the preceding run.

For both periodic and one-time jobs, you can also configure a job to analyze only those objects that are created or changed before or after a certain time or during a certain time range. To do this, add [object criteria](#) (p. 63) that use the last modified date for objects.

## Sampling depth

With this option, you specify the percentage of eligible S3 objects that you want a sensitive data discovery job to analyze. If this value is less than 100%, Macie selects eligible objects to analyze at random, up to the specified percentage, and analyzes all the data in those objects. For example, if you configure a job to analyze 10,000 objects and you specify a sampling depth of 20%, the job analyzes approximately 2,000 randomly selected, eligible objects.

Reducing the sampling depth of a job can lower the cost and reduce the duration of a job. It's helpful for cases where the data in objects is highly consistent and you want to determine whether an S3 bucket, rather than each object, contains sensitive data.

Note that this option controls the percentage of *objects* that are analyzed, not the percentage of *bytes* that are analyzed. If you enter a sampling depth that's less than 100%, Macie analyzes all the data in each selected object, not that percentage of the data in each selected object.

## S3 object criteria

To fine tune the scope of a sensitive data discovery job, you can also define custom criteria that determine which S3 objects are included or excluded from a job's analysis. These criteria consist of one or more conditions that derive from properties of S3 objects. The conditions apply to objects in all the S3 buckets that a job is configured to analyze. If a bucket contains multiple versions of an object, the conditions apply to the latest version of the object.

If you define multiple conditions as object criteria, Macie uses AND logic to join the conditions. In addition, exclude conditions take precedence over include conditions. For example, if you include objects that have the .pdf file name extension and exclude objects that are larger than 5 MB, the job analyzes any object that has the .pdf file name extension, unless the object is larger than 5 MB.

You can define conditions that derive from any of the following properties of S3 objects.

### File name extension

This correlates to the file name extension of an S3 object. You can use this type of condition to include or exclude objects based on file type. To do this for multiple types of files, enter the file name extension for each type and separate each entry with a comma—for example: **docx, pdf, xlsx**. If you enter multiple file name extensions as values for a condition, Macie uses OR logic to join the values.

Note that values are case sensitive. In addition, Macie doesn't support the use of partial values or wildcard characters in this type of condition.

For information about the types of files that Macie can analyze, see [Supported file and storage formats](#) (p. 89).

### Last modified

This correlates to the **Last modified** field in Amazon S3. In Amazon S3, this field stores the date and time when an S3 object was created or last changed, whichever is latest.

For a sensitive data discovery job, this condition can be a specific date, a specific date and time, or an exclusive time range:

- To analyze objects that were last modified after a certain date or date and time, enter the values in the **From** fields.
- To analyze objects that were last modified before a certain date or date and time, enter the values in the **To** fields.
- To analyze objects that were last modified during a certain time range, use the **From** fields to enter the values for the first date or date and time in the time range. Use the **To** fields to enter the values for the last date or date and time in the time range.
- To analyze objects that were last modified at any time during a certain single day, enter the date in the **From** date field. Enter the date for the next day in the **To** date field. Then verify that both time fields are blank. (Macie treats a blank time field as 00:00:00.) For example, to analyze objects that changed on August 9, 2020, enter **2020/08/09** in the **From** date field, enter **2020/08/10** in the **To** date field, and don't enter a value in either time field.

Enter any time values in Coordinated Universal Time (UTC) and use 24-hour notation.

### Prefix

This correlates to the **Key** field in Amazon S3. In Amazon S3, this field stores the name of an S3 object, including the object's prefix. A *prefix* is similar to a directory path within a bucket. It enables you to group similar objects together in a bucket, much like you might store similar files together in a folder on a file system. For information about object prefixes and folders in Amazon S3, see

[Organizing objects in the Amazon S3 console using folders](#) in the *Amazon Simple Storage Service User Guide*.

You can use this type of condition to include or exclude objects whose keys (names) begin with a certain value. For example, to exclude all objects whose key begins with *AWSLogs*, enter **AWSLogs** as the value for a **Prefix** condition, and then choose **Exclude**.

If you enter multiple prefixes as values for a condition, Macie uses OR logic to join the values. For example, if you enter **AWSLogs1** and **AWSLogs2** as values for a condition, any object whose key begins with *AWSLogs1* or *AWSLogs2* meets the condition's criteria.

When you enter a value for a **Prefix** condition, keep the following in mind:

- Values are case sensitive.
- Macie doesn't support the use of wildcard characters in these values.
- In Amazon S3, an object's key doesn't include the name of the bucket that contains the object. For this reason, don't specify bucket names in these values.
- If a prefix includes a delimiter, include the delimiter in the value. For example, enter **AWSLogs/eventlogs** to define a condition for all objects whose key begins with *AWSLogs/eventlogs*. Macie supports the default Amazon S3 delimiter, which is a slash (/), and custom delimiters.

Also note that an object meets a condition's criteria only if the object's key exactly matches the value that you enter, starting with the first character in the object's key. In addition, Macie applies a condition to the complete **Key** value for an object, including the object's file name.

For example, if an object's key is *AWSLogs/eventlogs/testlog.csv* and you enter any of the following values for a condition, the object meets the condition's criteria:

- **AWSLogs**
- **AWSLogs/event**
- **AWSLogs/eventlogs/**
- **AWSLogs/eventlogs/testlog**
- **AWSLogs/eventlogs/testlog.csv**

However, if you enter **eventlogs**, the object doesn't meet the criteria—the condition's value doesn't include the first part of the key, *AWSLogs/*. Similarly, if you enter **awslogs**, the object doesn't meet the criteria due to differences in capitalization.

### Storage size

This correlates to the **Size** field in Amazon S3. In Amazon S3, this field indicates the total storage size of an S3 object. If an object is a compressed file, this value doesn't reflect the actual size of the file after it's decompressed.

You can use this type of condition to include or exclude objects that are smaller than a certain size, larger than a certain size, or fall within a certain size range. Macie applies this type of condition to all types of objects, including compressed or archive files and the files that they contain. For information about size-based restrictions for each supported format, see [Amazon Macie quotas](#) (p. 210).

### Tags

Tags are custom labels that you can associate with specific types of AWS resources, including S3 objects. Each tag consists of a required tag key and an optional tag value. For information about tagging S3 objects, see [Categorizing your storage using tags](#) in the *Amazon Simple Storage Service User Guide*.

For a sensitive data discovery job, you can use this type of condition to include or exclude objects that have a specific tag. This can be a specific tag key or a specific tag key and tag value (as a pair).

If you specify multiple tags as values for a condition, Macie uses OR logic to join the values. For example, if you specify **Project1** and **Project2** as tag keys for a condition, any object that has the *Project1* or *Project2* tag key meets the condition's criteria.

Note that tag keys and values are case sensitive. In addition, Macie doesn't support use of partial values or wildcard characters in this type of condition.

## Creating a sensitive data discovery job

With Amazon Macie, you create and run sensitive data discovery jobs to automate discovery, logging, and reporting of sensitive data in Amazon Simple Storage Service (Amazon S3) buckets. A sensitive data discovery job analyzes objects in S3 buckets to determine whether the objects contain sensitive data, and it provides detailed reports of the sensitive data that it finds and the analysis that it performs.

When you create a job, you start by specifying which S3 buckets you want the job to analyze—specific buckets that you select or buckets that match specific criteria. Then you specify how often to run the job—once, or periodically on a daily, weekly, or monthly basis. You can also choose various options to refine the scope of the job's analysis. These options include custom criteria that derive from properties of S3 objects, such as last modified date and prefix.

After you define the schedule and scope of the job, you specify which managed data identifiers and custom data identifiers you want the job to use when it analyzes data:

- A *managed data identifier* is a set of built-in criteria and techniques that are designed to detect a specific type of sensitive data—for example, credit card numbers, AWS secret keys, or passport numbers for a particular country or region. These identifiers can detect a large and growing list of sensitive data types for many countries and regions, including multiple types of financial data, personal health information (PHI), and personally identifiable information (PII). For more information, see [Using managed data identifiers \(p. 37\)](#).
- A *custom data identifier* is a set of criteria that you define to detect sensitive data. With custom data identifiers, you can detect sensitive data that reflects your organization's particular scenarios, intellectual property, or proprietary data—for example, employee IDs, customer account numbers, or internal data classifications. These identifiers can supplement the managed data identifiers that Macie provides. For more information, see [Building custom data identifiers \(p. 53\)](#).

When you finish choosing these options, you're ready to specify a name for the job, and then review and save the job.

### Tasks

- [Before you begin \(p. 65\)](#)
- [Step 1: Choose S3 buckets \(p. 66\)](#)
- [Step 2: Review your S3 bucket selections or criteria \(p. 68\)](#)
- [Step 3: Define the schedule and refine the scope \(p. 68\)](#)
- [Step 4: Select managed data identifiers \(p. 69\)](#)
- [Step 5: Select custom data identifiers \(p. 70\)](#)
- [Step 6: Enter a name and description \(p. 70\)](#)
- [Step 7: Review and create \(p. 71\)](#)

## Before you begin

Before you create a job, it's a good idea to take the following steps:

- Verify that you configured Macie to store your sensitive data discovery results in an S3 bucket. To do this, choose **Discovery results** in the navigation pane on the Amazon Macie console, and then verify that you entered the settings. To learn about these settings, see [Storing and retaining sensitive data discovery results \(p. 96\)](#).
- Create any custom data identifiers that you want the job to use. To learn how, see [Building custom data identifiers \(p. 53\)](#).
- If you want the job to analyze objects that are encrypted with a customer managed AWS KMS key, ensure that Macie has permission to use the key. For more information, see [Analyzing encrypted S3 objects \(p. 90\)](#).
- If you want the job to analyze objects in a bucket that has a restrictive bucket policy, ensure that Macie is allowed to access objects in the bucket. For more information, see [Allowing Macie to access S3 buckets and objects \(p. 32\)](#).

If you do these things before you create a job, you streamline creation of the job and help ensure that the job analyzes the data that you want.

## Step 1: Choose S3 buckets

The first step in creating a job is to specify which S3 buckets you want the job to analyze. For this step, you have two options:

- **Select specific buckets** – With this option, you explicitly select each S3 bucket that you want the job to analyze. Then, when the job runs, it analyzes objects only in the buckets that you select.
- **Specify bucket criteria** – With this option, you define runtime criteria that determine which S3 buckets the job analyzes. The criteria consist of one or more conditions that derive from bucket properties. Then, when the job runs, it identifies buckets that match your criteria and analyzes objects in those buckets.

For detailed information about these options, see [Scope options for sensitive data discovery jobs \(p. 57\)](#).


The following sections provide step-by-step instructions for choosing and configuring each option. Choose the section for the option that you want.

### Select specific buckets

If you choose to explicitly select each S3 bucket that you want the job to analyze, Macie provides you with a complete inventory of your buckets in the current AWS Region. You can then use this inventory to select one or more buckets for the job to analyze. To learn about this inventory, see [Selecting S3 buckets \(p. 58\)](#).

If you're the Macie administrator for an organization, the inventory includes buckets that are owned by member accounts in your organization. You can configure the job to analyze objects in as many as 1,000 of these buckets, spanning as many as 1,000 accounts.

### To select specific buckets for the job

1. Open the Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Jobs**.
3. Choose **Create job**.
4. On the **Choose S3 buckets** page, choose **Select specific buckets**. Macie displays a table of all the buckets for your account in the current Region.
5. Under **Select S3 buckets**, optionally choose refresh () to retrieve the latest bucket metadata from Amazon S3.



If the information icon (i) appears next to any bucket names, we recommend that you do this. This icon indicates that a bucket was created during the past 24 hours, possibly after Macie last retrieved bucket and object metadata from Amazon S3 as part of the [daily refresh cycle](#) (p. 10).

6. In the table, select the check box for each bucket that you want the job to analyze.

#### Tip

- To find specific buckets more easily, enter filter criteria in the filter bar above the table. You can also sort the table by choosing a column heading.
- To quickly determine whether you already configured a job to periodically analyze objects in a bucket, refer to the **Monitored** column. If *Yes* appears in the column, the bucket is explicitly included in a periodic job or the bucket matched the criteria for a periodic job within the past 24 hours. In addition, the status of at least one of those jobs is not *Cancelled*. Macie updates this data on a daily basis.
- To quickly determine when you most recently ran a periodic or one-time job to analyze objects in a bucket, refer to the **Latest job run** column. For additional information about that job, refer to the bucket's details.
- To display a bucket's details, choose the bucket's name. In addition to job-related information, the details panel provides statistics and other information about the bucket, such as the bucket's public access settings. To learn more about this data, see [Reviewing your S3 bucket inventory](#) (p. 18).

7. When you finish selecting buckets, choose **Next**.

In the next step, you'll review and verify your selections.

### Specify bucket criteria

If you choose to specify runtime criteria that determine which S3 buckets the job analyzes, Macie provides options to help you choose fields, operators, and values for individual conditions in the criteria. To learn more about these options, see [Specifying S3 bucket criteria](#) (p. 60).

#### To specify bucket criteria for the job

1. Open the Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Jobs**.
3. Choose **Create job**.
4. On the **Choose S3 buckets** page, choose **Specify bucket criteria**.
5. Under **Specify bucket criteria**, do the following to add a condition to the criteria:
  - a. Place your cursor in the filter bar, and then choose the bucket property to use for the condition.
  - b. In the first field, choose an operator for the condition, **Equals** or **Not equals**.
  - c. In the next field, enter one or more values for the property.

Depending on the type and nature of the bucket property, Macie displays different options for entering values. For example, if you choose the **Effective permission** property, Macie displays a list of values to choose from. If you choose the **Account ID** property, Macie displays a text box in which you can enter one or more AWS account IDs. To enter multiple values in a text box, enter each value and separate each entry with a comma.

- d. Choose **Apply**. Macie adds the condition to a filter box below the filter bar.

By default, Macie adds the condition with an include statement. This means that the job is configured to analyze (*include*) objects in buckets that match the condition. To skip (*exclude*) buckets that match the condition, choose **Include** in the filter box, and then choose **Exclude**.



- e. Repeat the preceding steps for each additional condition that you want to add to the criteria.
6. To test your criteria, expand the **Preview the criteria results** section. This section displays a table of all the buckets that currently match the criteria.
7. To refine your criteria, do any of the following:
  - To remove a condition, choose **X** in the filter box for the condition.
  - To change a condition, remove the condition by choosing **X** in the filter box for the condition. Then add a condition that has the correct settings.
  - To remove all conditions, choose **Clear filters**.

Macie updates the table of criteria results to reflect your changes.

8. When you finish specifying bucket criteria, choose **Next**.

In the next step, you'll review and verify your criteria.

## Step 2: Review your S3 bucket selections or criteria

For this step, verify that you chose the correct settings in the preceding step.

### Review your bucket selections

If you selected specific S3 buckets for the job, review the table of buckets and change your bucket selections as necessary. The table provides insight into the projected scope and cost of the job's analysis. The data is based on the size and types of objects that are currently stored in a bucket.

The **Estimated cost** field indicates the total estimated cost (in US Dollars) of analyzing objects in a bucket. Each estimate reflects the projected amount of uncompressed data that the job will analyze in a bucket. If any objects are compressed or archive files, the estimate assumes that the files use a 3:1 compression ratio and the job can analyze all extracted files. For more information, see [Forecasting and monitoring costs for sensitive data discovery jobs \(p. 87\)](#).

### Review your bucket criteria

If you specified bucket criteria for the job, review each condition in the criteria. To change the criteria, choose **Previous**, and then use the filter settings in the preceding step to enter the correct criteria. When you finish, choose **Next**.

When you finish reviewing and verifying the settings, choose **Next**.

## Step 3: Define the schedule and refine the scope

For this step, specify how often you want the job to run—once, or periodically on a daily, weekly, or monthly basis. Also choose various options to refine the scope of the job's analysis. To learn about these options, see [Scope options for sensitive data discovery jobs \(p. 57\)](#).

### To define the schedule and refine the scope of the job

1. On the **Refine the scope** page, choose how often you want the job to run:
  - To run the job only once, immediately after you finish creating it, choose **One-time job**.
  - To run the job periodically on a recurring basis, choose **Scheduled job**. For **Update frequency**, choose whether to run the job daily, weekly, or monthly. Then use the **Include existing objects** option to define the scope of the job's first run:

- Select this check box to analyze all existing objects immediately after you finish creating the job. Each subsequent run analyzes only those objects that are created or changed after the preceding run.
- Clear this check box to skip analysis of all existing objects. The job's first run analyzes only those objects that are created or changed after you finish creating the job and before the first run starts. Each subsequent run analyzes only those objects that are created or changed after the preceding run.

Clearing this check box is helpful for cases where you've already analyzed the data and want to continue to analyze it periodically. For example, if you previously used Amazon Macie Classic to classify data and you recently moved to Macie, you might use this option to ensure continued discovery and classification of your data without incurring unnecessary costs or duplicating classification data.

2. (Optional) To specify the percentage of objects that you want the job to analyze, enter the percentage in the **Sampling depth** box. If this value is less than 100%, Macie selects the objects to analyze at random, up to the specified percentage, and analyzes all the data in those objects. The default value is 100%.
3. (Optional) To add specific criteria that determine which S3 objects are included or excluded from the job's analysis, expand the **Additional settings** section, and then enter the criteria. These criteria consist of individual conditions that derive from properties of objects.
  - To analyze (*include*) objects that meet a specific condition, enter the condition type and value, and then choose **Include**.
  - To skip (*exclude*) objects that meet a specific condition, enter the condition type and value, and then choose **Exclude**.

Repeat this step for each include or exclude condition that you want.

In Macie, exclude conditions take precedence over include conditions. For example, if you include objects that have the .pdf file name extension and exclude objects that are larger than 5 MB, the job analyzes any object that has the .pdf file name extension, unless the object is larger than 5 MB.

4. When you finish, choose **Next**.

## Step 4: Select managed data identifiers

For this step, specify which managed data identifiers you want the job to use when it analyzes S3 objects. You can configure the job to use all, some, or none of the managed data identifiers that Macie provides. To review a detailed list of the managed data identifiers that are currently available, see [Using managed data identifiers \(p. 37\)](#). We update that list each time we release a new managed data identifier.

If you choose to use only some managed data identifiers, Macie displays a table of the managed data identifiers that are currently available. You can use the table to select each managed data identifier that you want the job to use (*include*) or not use (*exclude*), depending on the selection type that you choose for the job. In the table, each managed data identifier's ID describes the type of sensitive data that the managed data identifier detects, for example: **USA\_PASSPORT\_NUMBER** for US passport numbers, **CREDIT\_CARD\_SECURITY\_CODE** for credit card verification codes, and **PGP\_PRIVATE\_KEY** for PGP private keys. To find specific identifiers more quickly, you can sort and filter the table by sensitive data category and type.

### To select managed data identifiers for the job

1. On the **Select managed data identifiers** page, under **Selection type**, do one of the following to specify which managed data identifiers you want the job to use:

- To use all managed data identifiers, choose **All**.

If you choose this option and you configured the job to run more than once, each run will automatically use new managed data identifiers that we release, in addition to all the managed data identifiers that are currently available.

- To exclude specific managed data identifiers, choose **Exclude**. Then, in the table that appears, select the check box for each managed data identifier that you don't want the job to use.

For example, if you don't want the job to detect and report occurrences of mailing addresses, select the **ADDRESS** check box. If you do this, the job will use all managed data identifiers except the one that detects mailing addresses.

If you choose the **Exclude** option and you configured the job to run more than once, each run will automatically use new managed data identifiers that we release, in addition to all the managed data identifiers that are currently available and you didn't explicitly exclude from the job.

- To include only specific managed data identifiers, choose **Include**. Then, in the table that appears, select the check box for each managed data identifier that you want the job to use.

For example, if you want the job to only detect and report occurrences of US passport numbers, select the **USA\_PASSPORT\_NUMBER** check box. If you do this, the job won't use any managed data identifiers except the one that detects US passport numbers.

- To exclude all managed data identifiers, choose **None**.

If you choose this option, the job won't use any managed data identifiers. In the [next step \(p. 70\)](#), configure the job to instead use one or more custom data identifiers that you specify.

2. When you finish, choose **Next**.


## Step 5: Select custom data identifiers

For this step, optionally select one or more [custom data identifiers \(p. 53\)](#) that you want the job to use when it analyzes S3 objects. The job will use the selected identifiers in addition to any managed data identifiers that you configured the job to use.

### To select custom data identifiers for the job

1. On the **Select custom data identifiers** page, select the check box for each custom data identifier that you want the job to use. You can select as many as 30 custom data identifiers.

#### Tip

To test or review the settings for a custom data identifier before you select it, choose the link icon () next to the identifier's name. Macie opens a page that displays the identifier's settings. You can also use this page to test the identifier with sample data. To do this, enter up to 1,000 characters of text in the **Sample data** box, and then choose **Submit**. Macie evaluates the sample data by using the identifier, and then reports the number of matches.

2. When you finish selecting custom data identifiers, choose **Next**.

## Step 6: Enter a name and description

For this step, specify a name and, optionally, a brief description of the job.

### To enter a name and description for the job

1. On the **Enter a name and description** page, enter a name for the job in the **Job name** box. The name can contain as many as 500 characters.

2. (Optional) For **Job description**, enter a brief description of the job. The description can contain as many as 200 characters.
3. When you finish, choose **Next**.

## Step 7: Review and create

For this final step, review the configuration settings for the job and verify that they're correct. This is an important step. After you create a job, you can't change any of its settings. This helps ensure that you have an immutable history of sensitive data findings and discovery results for data privacy and protection audits or investigations that you perform.

Depending on the job's settings, you can also review the total estimated cost (in US Dollars) of running the job once. If you selected specific S3 buckets for the job, the estimate is based on the size and types of objects in the buckets that you selected, and how much of that data the job can analyze. If you specified bucket criteria for the job, the estimate is based on the size and types of objects in as many as 500 buckets that currently match the criteria, and how much of that data the job can analyze. To learn about this estimate, see [Forecasting and monitoring costs for sensitive data discovery jobs \(p. 87\)](#).

### To review and create the job

1. On the **Review and create** page, review each setting and verify that it's correct. To change a setting, choose **Edit** in the section that contains the setting, and then enter the correct setting. You can also use the navigation tabs to go to the page that contains a setting.
2. When you finish verifying the settings, choose **Submit** to create and save the job. Macie checks the settings and notifies you of any issues to address.

#### Note

If you haven't configured a repository for your sensitive data discovery results, Macie displays a warning and doesn't save the job. To address this issue, choose **Configure** in the **Repository for sensitive data discovery results** section. Then enter the configuration settings for the repository. To learn how, see [Storing and retaining sensitive data discovery results \(p. 96\)](#). After you enter the settings, return to the **Review and create** page and refresh the **Repository for sensitive data discovery results** section of the page.

Although we don't recommend it, you can temporarily override the repository requirement and save the job. If you do this, you risk losing discovery results from the job—Macie will retain the results for only 90 days. To temporarily override the requirement, select the check box for the override option.

3. If Macie notifies you of issues to address, address the issues, and then choose **Submit** again to create and save the job.

If you configured the job to run once, on a daily basis, or on the current day of the week or month, Macie starts running the job immediately after you save it. Otherwise, Macie prepares to run the job on the specified day of the week or month. To monitor the job, you can [check the status of the job \(p. 85\)](#).

## Monitoring sensitive data discovery jobs with Amazon CloudWatch Logs

In addition to [monitoring the overall status \(p. 85\)](#) of a sensitive data discovery job, you can monitor and analyze specific events that occur as a job progresses. You can do this by using near-real-time logging data that Amazon Macie automatically publishes to Amazon CloudWatch Logs. The data in these logs provides a record of changes to a job's progress or status, such as the exact date and time when a job started to run, was paused, or finished running.

The log data also provides details about any account- or bucket-level errors that occur while a job runs. For example, if the permissions settings for an S3 bucket prevent a job from analyzing objects in the

bucket, Macie logs an event. The event indicates when the error occurred, and it identifies both the affected bucket and the account that owns the bucket. The data for these types of events can help you identify, investigate, and address errors that prevent Macie from analyzing the data that you want.

With Amazon CloudWatch Logs, you can monitor, store, and access log files from multiple systems, applications, and AWS services, including Macie. You can also query and analyze log data, and configure CloudWatch Logs to notify you when certain events occur or thresholds are met. CloudWatch Logs also provides features for archiving log data and exporting the data to Amazon S3. To learn more about CloudWatch Logs, see the [Amazon CloudWatch Logs User Guide](#).

#### Topics

- [How logging works for sensitive data discovery jobs \(p. 72\)](#)
- [Reviewing logs for sensitive data discovery jobs \(p. 73\)](#)
- [Log event schema for sensitive data discovery jobs \(p. 74\)](#)
- [Types of log events for sensitive data discovery jobs \(p. 75\)](#)

## How logging works for sensitive data discovery jobs

When you start running sensitive data discovery jobs, Macie automatically creates and configures the appropriate resources in Amazon CloudWatch Logs to log events for all of your jobs in the current AWS Region. Macie then publishes event data to those resources automatically when your jobs run. The permissions policy for the Macie [service-linked role \(p. 205\)](#) for your account allows Macie to perform these tasks on your behalf. You don't need to take any steps to create or configure resources in CloudWatch Logs, or to log event data for your jobs.

In CloudWatch Logs, logs are organized into *log groups*. Each log group contains *log streams*. Each log stream contains *log events*. The general purpose of each of these resources is as follows:

- A *log group* is a collection of log streams that share the same retention, monitoring, and access control settings—for example, the collection of logs for all of your sensitive data discovery jobs.
- A *log stream* is a sequence of log events that share the same source—for example, an individual sensitive data discovery job.
- A *log event* is a record of an activity that was recorded by an application or resource—for example, an individual event that Macie recorded and published for a particular sensitive data discovery job.

Macie publishes events for all of your sensitive data discovery jobs to one log group, and each job has a unique log stream in that log group. The log group has the following prefix and name:

```
/aws/macie/classificationjobs
```

If this log group already exists, Macie uses it to store log events for your jobs. This can be helpful if your organization uses automated configuration, such as [AWS CloudFormation](#), to create log groups with predefined log retention periods, encryption settings, tags, metric filters, and so on for job events.

If this log group doesn't exist, Macie creates it with the default settings that CloudWatch Logs uses for new log groups. The settings include a log retention period of **Never Expire**, which means that CloudWatch Logs stores the logs indefinitely. To change the retention period for the log group, you can use the Amazon CloudWatch console or the Amazon CloudWatch Logs API. To learn how, see [Working with log groups and log streams](#) in the *Amazon CloudWatch Logs User Guide*.

Within this log group, Macie creates a unique log stream for each job that you run, the first time that the job runs. The name of the log stream is the unique identifier for the job, such as `85a55dc0fa6ed0be5939d0408example`, in the following format.

```
/aws/macie/classificationjobs/85a55dc0fa6ed0be5939d0408example
```

Each log stream contains all the log events that Macie recorded and published for the corresponding job. For periodic jobs, this includes events for all of the job's runs. If you delete the log stream for a periodic job, Macie creates the stream again the next time that the job runs. If you delete the log stream for a one-time job, you can't restore it.

Note that logging is enabled by default for all of your jobs. You can't disable it or otherwise prevent Macie from publishing job events to CloudWatch Logs. If you don't want to store the logs, you can reduce the retention period for the log group to as little as one day. At the end of the retention period, CloudWatch Logs automatically deletes expired event data from the log group.

## Reviewing logs for sensitive data discovery jobs

You can review the logs for your sensitive data discovery jobs by using the Amazon CloudWatch console or the Amazon CloudWatch Logs API. Both the console and the API provide features that are designed to help you review and analyze log data. You can use these features to work with log streams and events for your jobs as you would work with any other type of log data in CloudWatch Logs.

For example, you can search and filter aggregate data to identify specific types of events that occurred for all of your jobs during a specific time range. Or you can perform a targeted review of all the events that occurred for a particular job. CloudWatch Logs also provides options for monitoring log data, defining metric filters, and creating custom alarms.

### Tip

To navigate to the log events for a particular job by using the Amazon Macie console, do the following: On the **Jobs** page, choose the name of the job. At the top of the details panel, choose **Show results**, and then choose **Show CloudWatch logs**. Macie opens the Amazon CloudWatch console and displays a table of log events for the job.

### To review the logs for your jobs (Amazon CloudWatch console)

1. Open the CloudWatch console at <https://console.aws.amazon.com/cloudwatch/>.
2. By using the AWS Region selector in the upper-right corner of the page, select the Region in which you ran jobs that you want to review logs for.
3. In the navigation pane, choose **Logs**, and then choose **Log groups**.
4. On the **Log groups** page, choose the **/aws/macie/classificationjobs** log group. CloudWatch Logs displays a table of log streams for the jobs that you've run. There is one unique stream for each job. The name of each stream correlates to the unique identifier for a job.
5. Under **Log streams**, do one of the following:
  - To review the log events for a particular job, choose the log stream for the job. To find the stream more easily, enter the job's unique identifier in the filter bar above the table. After you choose the log stream, CloudWatch Logs displays a table of log events for the job.
  - To review log events for all of your jobs, choose **Search all**. CloudWatch Logs displays a table of log events for all of your jobs.
6. (Optional) In the filter bar above the table, enter terms, phrases, or values that specify characteristics of specific events to review. For more information, see [Search log data using filter patterns](#) in the *Amazon CloudWatch Logs User Guide*.
7. To review the details of a specific log event, choose the right arrow (➡) in the row for the event. CloudWatch Logs displays the event's details in JSON format.

As you familiarize yourself with the data in the log events, you can also perform tasks such as [creating metrics filters](#) that turn log data into numerical CloudWatch metrics, and [creating custom alarms](#) that

make it easier for you to identify and respond to specific log events. For more information, see the [Amazon CloudWatch Logs User Guide](#).

## Log event schema for sensitive data discovery jobs

Each log event for a sensitive data discovery job is a JSON object that conforms to the Amazon CloudWatch Logs event schema and contains a standard set of fields. Some types of events have additional fields that provide information that's particularly useful for that type of event. For example, events for account-level errors include the account ID of the affected AWS account. Events for bucket-level errors include the name of the affected S3 bucket. For a detailed list of job events that Macie publishes to CloudWatch Logs, see [Types of log events for jobs \(p. 75\)](#).

The following example shows the log event schema for sensitive data discovery jobs. In this example, the event reports that Macie wasn't able to analyze any objects in an S3 bucket because Amazon S3 denied access to the bucket.

```
{
  "adminAccountId": "123456789012",
  "jobId": "85a55dc0fa6ed0be5939d0408example",
  "eventType": "BUCKET_ACCESS_DENIED",
  "occurredAt": "2021-04-14T17:11:30.574809Z",
  "description": "Macie doesn't have permission to access the affected S3 bucket.",
  "jobName": "My_Macie_Job",
  "operation": "ListObjectsV2",
  "runDate": "2021-04-14T17:08:30.345809Z",
  "affectedAccount": "111122223333",
  "affectedResource": {
    "type": "S3_BUCKET_NAME",
    "value": "DOC-EXAMPLE-BUCKET"
  }
}
```

In the preceding example, Macie attempted to list the objects in the bucket by using the [ListObjectsV2](#) operation of the Amazon S3 API. When Macie sent the request to Amazon S3, Amazon S3 denied access to the bucket.

The following fields are common to all log events for sensitive data discovery jobs:

- **adminAccountId** – The unique identifier for the AWS account that created the job.
- **jobId** – The unique identifier for the job.
- **eventType** – The type of event that occurred. For complete lists of possible values and a description of each one, see [Types of log events for jobs \(p. 75\)](#).
- **occurredAt** – The date and time, in Coordinated Universal Time (UTC) and extended ISO 8601 format, when the event occurred.
- **description** – A brief description of the event.
- **jobName** – The custom name of the job.

Depending on the type and nature of an event, a log event can also contain the following fields:

- **affectedAccount** – The unique identifier for the AWS account that owns the affected resource.
- **affectedResource** – An array that provides details about the affected resource. In the array, the `type` field specifies a field that stores metadata about a resource. The `value` field specifies the value for the field (`type`).
- **operation** – The operation that Macie attempted to perform and caused the error.
- **runDate** – The date and time, in Coordinated Universal Time (UTC) and extended ISO 8601 format, when the applicable job or job run started.



## Types of log events for sensitive data discovery jobs

Macie publishes log events for three categories of events:

- Job status events, which record changes to the status or progress of a job or a job run.
- Account-level error events, which record errors that prevented Macie from analyzing Amazon S3 data for a specific AWS account.
- Bucket-level error events, which record errors that prevented Macie from analyzing data in a specific S3 bucket.

The topics in this section list and describe the types of events that Macie publishes for each category.

### Topics

- [Job status events \(p. 75\)](#)
- [Account-level error events \(p. 77\)](#)
- [Bucket-level error events \(p. 79\)](#)

### Job status events

A job status event records a change to the status or progress of a job or a job run. For periodic jobs, Macie logs and publishes these events for both the overall job and individual job runs. For information about determining the overall status of a job, see [Checking the status of sensitive data discovery jobs \(p. 85\)](#).

The following example uses sample data to show the structure and nature of the fields in a job status event. In this example, a `SCHEDULED_RUN_COMPLETED` event indicates that a scheduled run of a periodic job finished running. The run started on April 14, 2021, at 17:09:30 UTC, as indicated by the `runDate` field. The run finished on April 14, 2021, at 17:16:30 UTC, as indicated by the `occurredAt` field.

```
{
  "adminAccountId": "123456789012",
  "jobId": "ffad0e71455f38a4c7c220f3cexample",
  "eventType": "SCHEDULED_RUN_COMPLETED",
  "occurredAt": "2021-04-14T17:16:30.574809Z",
  "description": "The scheduled job run finished running.",
  "jobName": "My_Daily_Macie_Job",
  "runDate": "2021-04-14T17:09:30.574809Z"
}
```

The following table lists and describes the types of job status events that Macie logs and publishes to CloudWatch Logs. The **Event type** column indicates the name of each event as it appears in the `eventType` field of an event. The **Description** column provides a brief description of the event as it appears in the `description` field of an event. The **Additional information** provides information about the type of job that the event applies to. The table is sorted first by the general chronological order in which events might occur, and then in ascending alphabetical order by event type.

Event type	Description	Additional information
JOB_CREATED	The job was created.	Applies to one-time and periodic jobs.
ONE_TIME_JOB_STARTED	The job started running.	Applies only to one-time jobs.
SCHEDULED_RUN_STARTED	The scheduled job run started running.	Applies only to periodic jobs. To log the start of a one-



Event type	Description	Additional information
		time job, Macie publishes a <code>ONE_TIME_JOB_STARTED</code> event, not this type of event.
<code>BUCKET_MATCHED_THE_CRITERIA</code>	The affected bucket matched the bucket criteria specified for the job.	Applies to one-time and periodic jobs that use runtime bucket criteria to determine which S3 buckets to analyze.  The <code>affectedResource</code> array specifies the name of the bucket that matched the criteria and was included in the job's analysis.
<code>NO_BUCKETS_MATCHED_THE_CRITERIA</code>	The job started running but no buckets currently match the bucket criteria specified for the job. The job didn't analyze any data.	Applies to one-time and periodic jobs that use runtime bucket criteria to determine which S3 buckets to analyze.
<code>SCHEDULED_RUN_COMPLETED</code>	The scheduled job run finished running.	Applies only to periodic jobs. To log completion of a one-time job, Macie publishes a <code>JOB_COMPLETED</code> event, not this type of event.
<code>JOB_PAUSED_BY_USER</code>	The job was paused by a user.	Applies to one-time and periodic jobs that you stopped temporarily (paused).
<code>JOB_RESUMED_BY_USER</code>	The job was resumed by a user.	Applies to one-time and periodic jobs that you stopped temporarily (paused) and subsequently resumed.
<code>JOB_PAUSED_BY_MACIE_SERVICE_QUOTA_WAS</code>	The job was paused by Macie. Completion of the job would exceed a monthly quota for the affected account.	Applies to one-time and periodic jobs that Macie stopped temporarily (paused).  Macie automatically pauses a job if completion of the job or a job run would exceed the monthly <a href="#">sensitive data discovery quota</a> (p. 210) for any accounts that the job analyzes data for. To avoid this issue, consider increasing the quota for the affected accounts.

Event type	Description	Additional information
JOB_RESUMED_BY_MACIE_SERVICE_QUOTA_LIMIT	The job was resumed by Macie. The monthly service quota was lifted for the affected account.	Applies to one-time and periodic jobs that Macie stopped temporarily (paused) and subsequently resumed.  If Macie automatically pauses a job because completion of the job or a job run would exceed the monthly <a href="#">sensitive data discovery quota</a> (p. 210) for an account, Macie automatically resumes the job when the subsequent month starts or the quota is increased for all the affected accounts.
JOB_CANCELLED	The job was cancelled.	Applies to one-time and periodic jobs that you stopped permanently (cancelled) or, for one-time jobs, paused and didn't resume within 30 days.  If you suspend or disable Macie, this type of event also applies to jobs that were active or paused when you suspended or disabled Macie. Macie automatically cancels your jobs in an AWS Region if you suspend or disable Macie in the Region.
JOB_COMPLETED	The job finished running.	Applies only to one-time jobs. To log completion of a job run for a periodic job, Macie publishes a SCHEDULED_RUN_COMPLETED event, not this type of event.

## Account-level error events

An account-level error event records an error that prevented Macie from analyzing objects in S3 buckets that are owned by a specific AWS account. The `affectedAccount` field in each event specifies the account ID for that account.

The following example uses sample data to show the structure and nature of the fields in an account-level error event. In this example, an `ACCOUNT_ACCESS_DENIED` event indicates that Macie wasn't able to analyze objects in any S3 buckets that are owned by account 444455556666.

```
{
  "adminAccountId": "123456789012",
  "jobId": "85a55dc0fa6ed0be5939d0408example",
  "eventType": "ACCOUNT_ACCESS_DENIED",
  "occurredAt": "2021-04-14T17:08:30.585709Z",
  "description": "Macie doesn't have permission to access S3 bucket data for the affected account.",
  "jobName": "My_Macie_Job",
  "operation": "ListBuckets",
```

```

    "runDate": "2021-04-14T17:05:27.574809Z",
    "affectedAccount": "444455556666"
  }

```

The following table lists and describes the types of account-level error events that Macie logs and publishes to CloudWatch Logs. The **Event type** column indicates the name of each event as it appears in the `eventType` field of an event. The **Description** column provides a brief description of the event as it appears in the `description` field of an event. The **Additional information** column provides any applicable tips for investigating or addressing the error that occurred. The table is sorted in ascending alphabetical order by event type.

Event type	Description	Additional information
ACCOUNT_ACCESS_DENIED	Macie doesn't have permission to access S3 bucket data for the affected account.	<p>This typically occurs because the buckets that are owned by the account have restrictive bucket policies. For information about how to address this issue, see <a href="#">Allowing Macie to access S3 buckets and objects (p. 32)</a>.</p> <p>The value for the <code>operation</code> field in the event can help you determine which permissions settings prevented Macie from accessing S3 data for the account. This field indicates the Amazon S3 operation that Macie attempted to perform when the error occurred.</p>
ACCOUNT_DISABLED	The job skipped resources that are owned by the affected account. Macie was disabled for the account.	To address this issue, re-enable Macie for the account in the same AWS Region.
ACCOUNT_DISASSOCIATED	The job skipped resources that are owned by the affected account. The account isn't associated with your Macie administrator account as a member account anymore.	<p>This occurs if you, as a Macie administrator for an organization, configure a job to analyze data for an associated member account and the member account is subsequently removed from your organization.</p> <p>To address this issue, re-associate the affected account with your Macie administrator account as a member account. For more information, see <a href="#">Managing multiple accounts (p. 184)</a>.</p>
ACCOUNT_ISOLATED	The job skipped resources that are owned by the affected account. The AWS account was isolated.	–

Event type	Description	Additional information
ACCOUNT_REGION_DISABLED	The job skipped resources that are owned by the affected account. The AWS account isn't active in the current AWS Region.	–
ACCOUNT_SUSPENDED	The job was cancelled or skipped resources that are owned by the affected account. Macie was suspended for the account.	<p>If the specified account is your own account, Macie automatically cancelled the job when you suspended Macie in the same Region. To address the issue, re-enable Macie in the Region.</p> <p>If the specified account is a member account, re-enable Macie for that account in the same Region.</p>
ACCOUNT_TERMINATED	The job skipped resources that are owned by the affected account. The AWS account was terminated.	–

## Bucket-level error events

A bucket-level error event records an error that prevented Macie from analyzing objects in a specific S3 bucket. The `affectedResource` array in each event specifies the name of the bucket. The `affectedAccount` field in each event specifies the account ID for the AWS account that owns the bucket.

The following example uses sample data to show the structure and nature of the fields in a bucket-level error event. In this example, a `BUCKET_ACCESS_DENIED` event indicates that Macie wasn't able to analyze any objects in the S3 bucket named `DOC-EXAMPLE-BUCKET`. When Macie attempted to list the objects in the bucket by using the [ListObjectsV2](#) operation of the Amazon S3 API, Amazon S3 denied access to the bucket.

```
{
  "adminAccountId": "123456789012",
  "jobId": "85a55dc0fa6ed0be5939d0408example",
  "eventType": "BUCKET_ACCESS_DENIED",
  "occurredAt": "2021-04-14T17:11:30.574809Z",
  "description": "Macie doesn't have permission to access the affected S3 bucket.",
  "jobName": "My_Macie_Job",
  "operation": "ListObjectsV2",
  "runDate": "2021-04-14T17:09:30.685209Z",
  "affectedAccount": "111122223333",
  "affectedResource": {
    "type": "S3_BUCKET_NAME",
    "value": "DOC-EXAMPLE-BUCKET"
  }
}
```

The following table lists and describes the types of bucket-level error events that Macie logs and publishes to CloudWatch Logs. The **Event type** column indicates the name of each event as it appears in the `eventType` field of an event. The **Description** column provides a brief description of the event

as it appears in the `description` field of an event. The **Additional information** column provides any applicable tips for investigating or addressing the error that occurred. The table is sorted in ascending alphabetical order by event type.

Event type	Description	Additional information
BUCKET_ACCESS_DENIED	Macie doesn't have permission to access the affected S3 bucket.	The value for the <code>operation</code> field in the event can help you determine which permissions settings prevented Macie from accessing the bucket. This field indicates the Amazon S3 operation that Macie attempted to perform when the error occurred.
BUCKET_DOES_NOT_EXIST	The affected S3 bucket doesn't exist anymore.	This typically occurs because a bucket was deleted.
BUCKET_IN_DIFFERENT_REGION	The affected S3 bucket was moved to a different AWS Region.	–
BUCKET_OWNER_CHANGED	The owner of the affected S3 bucket changed. Macie doesn't have permission to access the bucket anymore.	This typically occurs if ownership of a bucket was transferred to an AWS account that isn't part of your organization. The <code>affectedAccount</code> field in the event indicates the account ID for the account that previously owned the bucket.

## Reviewing statistics and results for a sensitive data discovery job

When you run a sensitive data discovery job, Amazon Macie automatically calculates and reports certain statistical data for the job. For example, Macie reports the number of times that the job has run and the approximate number of S3 objects that the job has yet to process during its current run.

As a job progresses, Macie also produces several types of results for the job: log events, sensitive data findings, and sensitive data discovery results.

### Log event

This is a record of an event that occurred while the job was running. Macie automatically logs and publishes data for certain events to Amazon CloudWatch Logs. The data in these logs provides a record of changes to the job's progress or status, such as the exact date and time when the job started or stopped running. The data also provides details about any account- or bucket-level errors that occurred while the job ran.

Log events can help you monitor a job and address any issues that prevented the job from analyzing the data that you want. If a job uses runtime criteria to determine which S3 buckets to analyze, log events can also help you determine whether and which S3 buckets matched the criteria when the job ran.

You can access log events by using the Amazon CloudWatch console or the Amazon CloudWatch Logs API. To help you navigate to the log events for a job, the Amazon Macie console provides a link to them. For more information, see [Monitoring jobs \(p. 71\)](#).

### **Sensitive data finding**

This is a report of sensitive data that Macie found in an object. Each finding provides a severity rating and details such as:

- The date and time when Macie found the sensitive data.
- The category and types of sensitive data that Macie found.
- The number of occurrences of each type of sensitive data that Macie found.
- The location of as many as 15 occurrences of the sensitive data that Macie found.
- The unique identifier for the job that produced the finding.
- The name, public access settings, encryption type, and other information about the affected S3 bucket and object.

A sensitive data finding doesn't include the sensitive data that Macie found. Instead, it provides information that you can use for further investigation and remediation as necessary.

Macie stores sensitive data findings for 90 days. You can access them by using the Amazon Macie console or the Amazon Macie API. You can also monitor and process them by using other applications, services, and systems. For more information, see [Analyzing findings \(p. 104\)](#).

### **Sensitive data discovery result**

This is a record that logs details about the analysis of an object. Macie creates a sensitive data discovery result for each object that you configure a job to analyze. This includes objects that don't contain sensitive data, and therefore don't produce a sensitive data finding, and objects that Macie can't analyze due to issues such as permissions settings or use of an unsupported format.

If an object does contain sensitive data, the sensitive data discovery result includes data from the corresponding sensitive data finding. It provides additional information too, such as the location of as many as 1,000 occurrences of each type of sensitive data that Macie found in the object. For example:

- The column and row number for a cell or field in a Microsoft Excel workbook, CSV file, or TSV file
- The path to a field or array in a JSON or JSON Lines file
- The line number for a line in a non-binary text file other than a CSV, JSON, JSON Lines, or TSV file—for example, an HTML, TXT, or XML file
- The page number for a page in an Adobe Portable Document Format (PDF) file
- The record index and the path to a field in a record in an Apache Avro object container or Apache Parquet file

Note that a sensitive data discovery result doesn't include the sensitive data that Macie found. Instead, it provides you with an analysis record that can be helpful for data privacy and protection audits or investigations.

Macie stores sensitive data discovery results for 90 days. You can't access them directly on the Amazon Macie console or through the Amazon Macie API. Instead, you configure Macie to store the results in an S3 bucket, and then optionally access and query the results in that bucket. This configuration also ensures long-term storage and retention of the results. To learn how to configure these settings, see [Storing and retaining sensitive data discovery results \(p. 96\)](#).

After you configure Macie to store your discovery results in an S3 bucket, Macie writes the results to JSON Lines (.jsonl) files and adds those files to the bucket as GNU Zip (.gz) files. To help you navigate to the results, the Amazon Macie console provides links to them.

Sensitive data findings and sensitive data discovery results both adhere to standardized schemas. This can help you optionally query, monitor, and process them by using other applications, services, and systems.

### To review statistics and results for a job

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Jobs**.
3. On the **Jobs** page, choose the name of the job whose statistics and results you want to review. The details panel displays statistics, settings, and other information about the job.
4. In the details panel, do any of the following:
  - To review processing statistics for the job, refer to the **Statistics** section of the panel. This section displays statistics such as the number of times that the job has run and the approximate number of objects that the job has yet to process during its current run.
  - To review log events for the job, choose **Show results** at the top of the panel, and then choose **Show CloudWatch logs**. Macie opens the Amazon CloudWatch console and displays a table of the log events that Macie published for the job.
  - To review all the sensitive data findings that the job produced, choose **Show results** at the top of the panel, and then choose **Show findings**. Macie opens the **Findings** page and displays all the findings from the job. To review the details of a particular finding, choose the finding in the table and refer to the details panel.

#### Tip

In the finding details panel, you can use the link in the **Detailed result location** field to navigate to a finding's corresponding sensitive data discovery result in Amazon S3:

- If the finding applies to a large archive or compressed file, the link displays the folder that contains the discovery results for the file. An archive or compressed file is *large* if it generates more than 100 discovery results.
- If the finding applies to a small archive or compressed file, the link displays the file that contains the discovery results for the file. An archive or compressed file is *small* if it generates 100 or fewer discovery results.
- If the finding applies to another type of file, the link displays the file that contains the discovery results for the file.
- To review all the sensitive data discovery results that the job produced, choose **Show results** at the top of the panel, and then choose **Show classifications**. Macie opens the Amazon S3 console and displays the folder that contains all the discovery results for the job. This option is available only after you configure Macie to [store your sensitive data discovery results \(p. 96\)](#) in an S3 bucket.

## Managing sensitive data discovery jobs

To help you manage your sensitive data discovery jobs, Amazon Macie provides a complete inventory of your jobs in each AWS Region. With this inventory, you can manage your jobs as a single collection, and access the configuration settings, status, and processing statistics for individual jobs. You can also access the [sensitive data findings and other results \(p. 80\)](#) that each job produced.

In addition to these tasks, you can create custom variations of individual jobs—copy an existing job, adjust the settings for the copy, and then save the copy as a new job. This can be helpful for cases where you want to analyze different sets of data in the same way, or the same set of data in different ways. Or you want to adjust the configuration settings for an existing job—cancel the existing job, copy it, and then adjust and save the copy as a new job.

### Topics

- [Reviewing your inventory of sensitive data discovery jobs \(p. 83\)](#)
- [Reviewing configuration settings for sensitive data discovery jobs \(p. 83\)](#)
- [Checking the status of sensitive data discovery jobs \(p. 85\)](#)
- [Pausing, resuming, or cancelling sensitive data discovery jobs \(p. 86\)](#)
- [Copying sensitive data discovery jobs \(p. 86\)](#)

## Reviewing your inventory of sensitive data discovery jobs

The **Jobs** page on the Amazon Macie console provides information about all the sensitive data discovery jobs for your account in the current AWS Region. For each job, the table displays summary information that includes: the current status of the job; whether the job runs on a scheduled, periodic basis; and whether the job analyzes a specific number of S3 buckets or it analyzes S3 buckets that match runtime criteria. If you choose a job in the table, the details panel displays the configuration settings and other information about the job.

### To review your job inventory

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Jobs**. The **Jobs** page opens and displays the number of jobs in your inventory and a table of those jobs.
3. To find a specific job more quickly, do any of the following:
  - To sort the table by a specific field, click the column heading for the field. To change the sort order, click the column heading again.
  - To show only those jobs that have a specific value for a field, place your cursor in the filter bar. In the menu that appears, choose the field to use for the filter, and enter the value for the filter. Then choose **Apply**.
  - To hide jobs that have a specific value for a field, place your cursor in the filter bar. In the menu that appears, choose the field to use for the filter, and enter the value for the filter. Then choose **Apply**. In the filter bar, choose the equals icon (●) in the filter box. This changes the filter's operator from *equals* to *not equals* (⊘).
  - To remove a filter, choose the remove filter icon (⊗) in the filter box for the filter to remove.
4. To review the configuration settings and other details for a particular job, choose the job's name in the table, and then refer to the details panel.

## Reviewing configuration settings for sensitive data discovery jobs

On the Amazon Macie console, you can use the details panel on the **Jobs** page to review configuration settings and other information about an individual sensitive data discovery job. For example, you can review a list of the S3 buckets that a job is configured to analyze and which managed data identifiers a job uses to analyze objects in those buckets.

Note that you can't change any settings for an existing job. This helps ensure that you have an immutable history of sensitive data findings and discovery results for data privacy and protection audits or investigations that you perform. If you want to change an existing job, you can [cancel the job \(p. 86\)](#). Then [copy the job \(p. 86\)](#), configure the copy to use the settings that you want, and save the copy as a new job. If you do this, you should take steps to ensure that the new job doesn't analyze existing data in the same way again. To do this, note the date and time when you cancel the existing job. Then configure the scope of the new job to include only those objects that are created or



changed after you cancel the original job. For example, use object criteria to add a **Last modified** exclude condition that specifies the date and time when you cancelled the original job. For more information, see [Scope options for jobs \(p. 57\)](#).

### To review a job's configuration settings

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Jobs**.
3. On the **Jobs** page, choose the name of the job whose settings you want to review. The details panel displays the configuration settings and other information about the job.

Depending on the job's settings, the panel contains the following sections:

- **General information** – This section indicates the current status of the job and it provides general information about the job—for example, the Amazon Resource Name (ARN) of the job and the most recent date and time when the job started to run. If you paused the job during the past 30 days, this section also indicates when you paused the job and when the job or job run will expire if you don't resume it.
- **Statistics** – This section shows processing statistics for the job—for example, the number of times that the job has run and the approximate number of objects that the job has yet to process during its current run.
- **Scope** – This section indicates how often the job runs. It also shows the settings that refine the scope of the job—for example, the sampling depth and any [object criteria \(p. 63\)](#) that include or exclude S3 objects from the job's analysis.
- **S3 buckets** – This section appears in the panel if the job is configured to analyze buckets that you explicitly selected when you created the job. It indicates the number of AWS accounts that the job is configured to analyze data for. It also indicates the number of buckets that the job is configured to analyze and the names of those buckets (grouped by account). To show the complete list of accounts and buckets in JSON format, choose the number in the **Total buckets** field.
- **S3 bucket criteria** – This section appears in the panel if the job uses runtime criteria to determine which buckets to analyze. It lists any inclusion and exclusion criteria that the job is configured to use. To review the criteria in JSON format, choose **Details**, and then choose the **Criteria** tab in the window that appears.

#### Tip

To review a table of buckets that currently match the criteria, choose **Details**, and then choose the **Matching buckets** tab in the window that appears. Optionally choose refresh



to retrieve the latest data.

If the job has already run, you can also determine whether any buckets matched the criteria when the job ran and, if so, the names of those buckets. You can do this by reviewing the [job status log events \(p. 71\)](#) for the job. To do this, choose **Show results** at the top of the panel, and then choose **Show CloudWatch logs**. Macie opens the Amazon CloudWatch console and displays a table of log events for the job, including a `BUCKET_MATCHED_THE_CRITERIA` event for each bucket that matched the criteria and was included in the job's analysis.

- **Managed data identifiers** – This section indicates which [managed data identifiers \(p. 37\)](#) the job is configured to use when it analyzes objects. This is determined by the managed data identifier selection type for the job:
  - **Include all** – Use all the managed data identifiers that are available when the job runs.
  - **Include selected** – Use only the managed data identifiers listed in the **Selections** section.
  - **Exclude selected** – Use all the managed data identifiers that are available when the job runs, except the ones listed in the **Selections** section.
  - **Exclude all** – Don't use any managed data identifiers.

To review these settings in JSON format, choose **Details**.

- **Custom data identifiers** – This section appears in the panel if the job is configured to use custom data identifiers when it analyzes objects. It lists the names of those custom data identifiers.
4. (Optional) To review and save the job's settings in JSON format, choose the unique identifier for the job (**Job ID**) at the top of the panel, and then choose **Download**.

## Checking the status of sensitive data discovery jobs

When you create a sensitive data discovery job, its initial status is **Active (Running)** or **Active (Idle)**, depending on the job's type and schedule. The job then passes through additional states, which you can monitor as the job progresses.

### Tip

In addition to monitoring the overall status of a job, you can monitor specific events that occur as a job progresses. You can do this by using logging data that Macie automatically publishes to Amazon CloudWatch Logs. The data in these logs provides a record of changes to a job's status and details about any account- or bucket-level errors that occur while a job runs. For more information, see [Monitoring jobs \(p. 71\)](#).

### To check the status of a job

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Jobs**.
3. On the **Jobs** page, locate the job whose status you want to check. The **Status** field indicates the current status of the job:
  - **Active (Idle)** – For a periodic job, the previous run is complete and the next scheduled run is pending. This value doesn't apply to one-time jobs.
  - **Active (Running)** – For a one-time job, the job is currently in progress. For a periodic job, a scheduled run is in progress.
  - **Cancelled** – For any type of job, the job was stopped permanently (cancelled). A job has this status if you explicitly cancelled it or, if it's a one-time job, you paused the job and didn't resume it within 30 days. A job can also have this status if you [suspended Macie \(p. 200\)](#) in the current AWS Region.
  - **Complete** – For a one-time job, the job ran successfully and is now complete. This value doesn't apply to periodic jobs. Instead, the status of a periodic job changes to **Active (Idle)** when each run completes successfully.
  - **Paused (By Macie)** – For any type of job, the job was stopped temporarily (paused) by Macie. A job has this status if completion of the job or a job run would exceed the monthly [sensitive data discovery quota \(p. 210\)](#) for your account or any member accounts that the job analyzes data for. When this happens, Macie automatically pauses the job. Macie automatically resumes the job when the subsequent month starts or the quota is increased for all the affected accounts.
  - **Paused (By user)** – For any type of job, the job was stopped temporarily (paused) by you.

If you pause a one-time job and you don't resume it within 30 days, the job expires and Macie cancels it. If you pause a periodic job while it's actively running and you don't resume it within 30 days, the job's run expires and Macie cancels the run. To check the expiration date for a paused job or job run, choose the job's name in the table, and then refer to the **Expires** field in the **Status details** section of the details panel.

If a job is cancelled or paused, you can refer to the job's details to determine whether the job started to run or, for a periodic job, ran at least once before it was cancelled or paused. To do this, choose the job's name in the table, and then refer to the details panel. In the panel, the **Number of runs** field indicates the number of times that the job has run. The **Last run time** field indicates the most recent date and time when the job started to run.

Depending on the job's current status, you can optionally pause, resume, or cancel the job.

## Pausing, resuming, or cancelling sensitive data discovery jobs

After you create a sensitive data discovery job, you can pause it temporarily or cancel it permanently. When you pause a job that's actively running, Macie immediately begins to pause all processing tasks for the job. When you cancel a job that's actively running, Macie immediately begins to stop all processing tasks for the job. You can't resume or restart a job after it's cancelled.

If you pause a one-time job, you can resume it within 30 days. When you resume the job, Macie immediately resumes processing from the point where you paused the job—Macie doesn't restart the job from the beginning. If you don't resume a one-time job within 30 days of pausing it, the job expires and Macie cancels it.

If you pause a periodic job, you can resume it at any time. If you resume a periodic job and the job was idle when you paused it, Macie resumes the job according to the schedule and other configuration settings that you chose when you created the job. If you resume a periodic job and the job was actively running when you paused it, how Macie resumes the job depends on when you resume the job:

- If you resume the job within 30 days of pausing it, Macie immediately resumes the latest scheduled run from the point where you paused the job—Macie doesn't restart the run from the beginning.
- If you don't resume the job within 30 days of pausing it, the latest scheduled run expires and Macie cancels all remaining processing tasks for the run. When you subsequently resume the job, Macie resumes the job according to the schedule and other configuration settings that you chose when you created the job.

To help you determine when a paused job or job run will expire, Macie adds an expiration date to the job's details while the job is paused. To check this date, choose the job's name in the table on the **Jobs** page, and then refer to the **Expires** field in the **Status details** section of the details panel. In addition, we notify you approximately seven days before the job or job run will expire. We notify you again when the job or job run expires and is cancelled. To notify you, we send email to the address that's associated with your AWS account. We also create AWS Health events and Amazon CloudWatch Events for your account.

### To pause, resume, or cancel a job

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Jobs**.
3. On the **Jobs** page, select the check box for the job that you want to pause, resume, or cancel, and then do one of the following on the **Actions** menu:
  - To pause the job temporarily, choose **Pause**. This option is available only if the job's current status is **Active (Idle)**, **Active (Running)**, or **Paused (By Macie)**.
  - To resume the job, choose **Resume**. This option is available only if the job's current status is **Paused (By user)**.
  - To cancel the job permanently, choose **Cancel**. If you choose this option, you can't subsequently resume or restart the job.

## Copying sensitive data discovery jobs

To quickly create a new sensitive data discovery job that's similar to an existing job, you can create a copy of the job, edit the copy's settings, and then save the copy as a new job. This can be helpful for cases where you want to create a custom variation of an existing job. Or you want to adjust the configuration settings for an existing job by cancelling the job, and then copying, changing, and saving the settings as a new job.

### To copy a job

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Jobs**.
3. Select the check box for the job that you want to copy.
4. On the **Actions** menu, choose **Copy to new**.
5. Complete the steps on the console to review and adjust the settings for the copy of the job. On the **Scope** page, consider choosing options that prevent the job from analyzing existing data in the same way again:
  - For a one-time job, use [object criteria \(p. 63\)](#) to include only those objects that were created or changed after a certain time. For example, if you're creating a copy of a job that you cancelled, add a **Last modified** condition that specifies the date and time when you cancelled the existing job.
  - For a periodic job, clear the **Include existing objects** check box. If you do this, the first run of the job analyzes only those objects that are created or changed after you create the job and before the job's first run. You can also use [object criteria \(p. 63\)](#) to exclude objects that were last modified before a certain date and time.
6. When you finish, choose **Submit** to save the copy as a new job.

## Forecasting and monitoring costs for sensitive data discovery jobs

Amazon Macie pricing is based partly on the amount of data that you analyze by running sensitive data discovery jobs. To forecast and monitor your estimated costs for running sensitive data discovery jobs, you can review cost estimates that Macie provides when you create a job and after you start running jobs.

To view and monitor your actual costs, you can use AWS Billing and Cost Management. AWS Billing and Cost Management provides features that are designed to help you track and analyze your costs for AWS services, and manage budgets for your account or organization. It also provides features that can help you forecast usage costs based on historical data. To learn more, see the [AWS Billing and Cost Management User Guide](#).

For information about Macie pricing, see [Amazon Macie pricing](#).

### Topics

- [Forecasting the cost of a sensitive data discovery job \(p. 87\)](#)
- [Monitoring estimated costs for sensitive data discovery jobs \(p. 89\)](#)

## Forecasting the cost of a sensitive data discovery job

When you create a sensitive data discovery job, Macie can calculate and display estimated costs during two key steps in the job creation process: when you review the table of S3 buckets that you selected for the job (step 2) and when you review all the settings for the job (step 6). These estimates can help you determine whether to adjust the job's settings before you save the job. The availability and nature of the estimates depends on the settings that you choose for the job.

### Reviewing estimated costs for individual buckets (step 2)

If you explicitly select individual buckets for a job to analyze, you can review the estimated cost of analyzing objects in each of those buckets. Macie displays these estimates during step 2 of the job creation process, when you review your bucket selections. In the **Select S3 buckets** table for this

step, the **Estimated cost** field indicates the total estimated cost (in US Dollars) of running the job once to analyze objects in a bucket.

Each estimate reflects the projected amount of uncompressed data that the job will analyze in a bucket, based on the size and types of objects that are currently stored in the bucket. The estimate also reflects Macie pricing for the current AWS Region.

Only classifiable objects are included in the cost estimate for a bucket. A *classifiable object* is an S3 object that uses a supported Amazon S3 storage class (S3 Intelligent-Tiering, S3 One Zone-IA, S3 Standard, or S3 Standard-IA) and has a file name extension for a [supported file or storage format](#) (p. 89). If any classifiable objects are compressed or archive files, the estimate assumes that the files use a 3:1 compression ratio and the job can analyze all extracted files.

### Reviewing the total estimated cost of a job (step 6)

If you create a one-time job or you create and configure a periodic job to include existing S3 objects, Macie calculates and displays the job's total estimated cost during the final step (step 6) of the job creation process. You can review this estimate while you review and verify all the settings that you selected for the job.

This estimate indicates the total projected cost (in US Dollars) of running the job once in the current Region. The estimate reflects the projected amount of uncompressed data that the job will analyze. It's based on the size and types of objects that are currently stored in buckets that you explicitly selected for the job or up to 500 buckets that currently match bucket criteria that you specified for the job, depending on the job's settings.

Note that this estimate doesn't reflect any options that you selected to refine and reduce the scope of the job—for example, a lower sampling depth, or criteria that exclude certain S3 objects from the job. It also doesn't reflect your monthly [sensitive data discovery quota](#) (p. 210), which might limit the scope and cost of the job's analysis, or any discounts that might apply to your account.

In addition to the total estimated cost of the job, the estimate provides aggregated data that offers insight into the projected scope and cost of the job:

- **Size** values indicate the total storage size of the objects that the job can and can't analyze.
- **Object count** values indicate the total number of objects that the job can and can't analyze.

In these values, a **Classifiable** object is an S3 object that uses a supported Amazon S3 storage class (S3 Intelligent-Tiering, S3 One Zone-IA, S3 Standard, or S3 Standard-IA) and has a file name extension for a [supported file or storage format](#) (p. 89). Only classifiable objects are included in the cost estimate. A **Not classifiable** object is an object that doesn't use a supported Amazon S3 storage class or doesn't have a file name extension for a supported file or storage format. These objects aren't included in the cost estimate.

The estimate provides additional aggregated data for S3 objects that are compressed or archive files. The **Compressed** value indicates the total storage size of objects that use a supported Amazon S3 storage class and have a file name extension for a supported type of compressed or archive file. The **Uncompressed** value indicates the approximate size of these objects if they're decompressed, based on a specified compression ratio. This data is relevant due to the way that Macie analyzes compressed files and archive files.

When Macie analyzes a compressed or archive file, it inspects both the full file and the contents of the file. To inspect the file's contents, Macie decompresses the file, and then inspects each extracted file that uses a supported format. The actual amount of data that a job analyzes therefore depends on:

- Whether a file uses compression and, if so, the compression ratio that it uses.
- The number, size, and format of the extracted files.

By default, Macie assumes the following when it calculates cost estimates for a job:

- All compressed and archive files use a 3:1 compression ratio.
- All the extracted files use a supported file or storage format.

These assumptions can result in a larger size estimate for the scope of the data that the job will analyze, and, consequently, a higher cost estimate for the job.

You can recalculate the job's total estimated cost based on a different compression ratio. To do this, choose the ratio from the **Choose an estimated compression ratio** list in the **Estimated cost** section. Macie then updates the estimate to match your selection.

For more information about how Macie calculates estimated costs, see [Forecasting and monitoring Amazon Macie costs \(p. 193\)](#).

## Monitoring estimated costs for sensitive data discovery jobs

If you're already running sensitive data discovery jobs, the **Usage** page on the Amazon Macie console can help you monitor the estimated cost of those jobs. The page shows your estimated costs (in US Dollars) for using Macie in the current AWS Region during the current calendar month. For information about how Macie calculates these estimates, see [Forecasting and monitoring Amazon Macie costs \(p. 193\)](#).

### To review your estimated costs for running jobs

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. By using the AWS Region selector in the upper-right corner of the page, select the Region in which you want to review your estimated costs.
3. In the navigation pane, choose **Usage**.
4. In the **Estimated costs** section, refer to the breakdown of estimated costs for your account. The **Data discovery jobs** item reports the total estimated cost of the jobs that you've run thus far during the current month in the current Region.

If you're the Macie administrator for an organization, the **Estimated costs** section shows estimated costs for your organization overall for the current month in the current Region. To show the total estimated cost of the jobs that were run for a specific account, choose the account in the table. The **Estimated costs** section then shows a breakdown of estimated costs for the account, including the estimated cost of the jobs that were run. To show this data for a different account, choose the account in the table. To clear your account selection, choose **X** next to the account ID.

To view and monitor your actual costs, use [AWS Billing and Cost Management](#).

## Supported file and storage formats in Amazon Macie

When Amazon Macie analyzes data, it performs a deep inspection that factors the file or storage format for the data. Macie can analyze data in many different formats, including commonly used compression and archive formats. This support applies to the use of both managed data identifiers and custom data identifiers.

When Macie analyzes a compressed or archive file, it inspects both the full file and the contents of the file. To inspect the file's contents, it decompresses the file, and then inspects each extracted file that uses a supported format. Macie can do this for as many as 1,000,000 files and up to a nested depth of 10 levels.

The following table lists and describes the file and storage formats that Macie can analyze, organized by type. For each supported type, it also lists the applicable file name extensions.

File or storage type	Description	File name extensions
Big data	Apache Avro object containers and Apache Parquet files	.avro, .parquet
Compression or archive	GNU Zip compressed archives, TAR archives, and ZIP compressed archives	.gz, .gzip, .tar, .zip
Document	Adobe Portable Document Format files, Microsoft Excel workbooks, and Microsoft Word documents	.doc, .docx, .pdf, .xls, .xlsx
Text	Non-binary text files such as comma-separated values (CSV) files, Hypertext Markup Language (HTML) files, JavaScript Object Notation (JSON) files, JSON Lines files, plain-text documents, tab-separated values (TSV) files, and Extensible Markup Language (XML) files	.csv, .htm, .html, .json, .jsonl, .tsv, .txt, .xml, and others (depending on the type of non-binary text file)

Macie doesn't analyze data in images or audio, video, and other types of multimedia content.

For information about the quotas that apply to sensitive data discovery, see [Amazon Macie quotas \(p. 210\)](#).

## Analyzing encrypted S3 objects with Amazon Macie

When you enable Amazon Macie for your AWS account, Macie creates a [service-linked role \(p. 205\)](#) that grants Macie the permissions that it requires to call Amazon Simple Storage Service (Amazon S3) and other AWS services on your behalf. The permissions policy for this role (*AWSServiceRoleForAmazonMacie*) allows Macie to perform actions that include retrieving information about your S3 buckets and objects, and retrieving and analyzing objects in your S3 buckets. If your account is the Macie administrator account for an organization, the policy also allows Macie to perform these actions for member accounts in your organization.

If an S3 object is encrypted, the permissions policy for the service-linked role typically grants Macie the permissions that it requires to decrypt and inspect the object for sensitive data. However, this depends on the type of encryption that was used to encrypt the object. It can also depend on whether Macie is allowed to use the appropriate encryption key.

### Topics

- [Encryption options for S3 objects \(p. 91\)](#)
- [Allowing Macie to use a customer managed AWS KMS key \(p. 92\)](#)



## Encryption options for S3 objects

Amazon S3 supports multiple encryption options for S3 objects. For most of these options, Macie can decrypt and analyze an object by using the *AWSServiceRoleForAmazonMacie* service-linked role for your account. However, this depends on the type of encryption that was used to encrypt an object.

### Server-side encryption with Amazon S3 managed keys (SSE-S3)

If an object is encrypted using server-side encryption with an Amazon S3 managed key, Macie can decrypt and analyze the object.

To learn about this type of encryption, see [Protecting data using server-side encryption with Amazon S3 managed encryption keys](#) in the *Amazon Simple Storage Service User Guide*.

### Server-side encryption with AWS KMS keys (SSE-KMS)

If an object is encrypted using server-side encryption with an AWS managed AWS KMS key, Macie can decrypt and analyze the object.

If an object is encrypted using server-side encryption with a customer managed AWS KMS key, Macie can decrypt and analyze the object only if you [allow Macie to use the key \(p. 92\)](#). Otherwise, Macie can only store and report metadata for the object.

To learn about this type of encryption, see [Protecting data using server-side encryption with KMS keys stored in AWS Key Management Service](#) in the *Amazon Simple Storage Service User Guide*.

### Server-side encryption with customer-provided keys (SSE-C)

If an object is encrypted using server-side encryption with a customer-provided key, Macie can't decrypt and analyze the object. Macie can only store and report metadata for the object.

To learn about this type of encryption, see [Protecting data using server-side encryption with customer-provided encryption keys](#) in the *Amazon Simple Storage Service User Guide*.

### Client-side encryption

If an object is encrypted using client-side encryption, Macie can't decrypt and analyze the object. Macie can only store and report metadata for the object. For example, Macie can report the size of the object and the tags that are associated with the object.

To learn about this type of encryption in the context of Amazon S3, see [Protecting data using client-side encryption](#) in the *Amazon Simple Storage Service User Guide*.

You can [filter your bucket inventory \(p. 24\)](#) in Macie to determine which S3 buckets contain objects that use certain types of encryption. You can also determine which buckets use certain types of server-side encryption by default when storing new objects. The following table provides some example filters that you can apply to your bucket inventory to find this information.

To show buckets that...	Apply this filter...
Contain objects that use SSE-C encryption	<b>Object count by encryption</b> is <b>Customer managed</b> and <b>From</b> = 1
Contain objects that use SSE-KMS encryption	<b>Object count by encryption</b> is <b>SSE-KMS managed</b> and <b>From</b> = 1
Contain objects that use SSE-S3 encryption	<b>Object count by encryption</b> is <b>SSE-S3 managed</b> and <b>From</b> = 1



To show buckets that...	Apply this filter...
Contain objects that use client-side encryption (or aren't encrypted)	<b>Object count by encryption</b> is <b>No encryption</b> and <b>From</b> = 1
Encrypt new objects by default using SSE-KMS encryption	<b>Default encryption</b> = <b>aws:kms</b>
Encrypt new objects by default using SSE-S3 encryption	<b>Default encryption</b> = <b>AES256</b>

If a bucket is configured to encrypt new objects by default using SSE-KMS encryption, you can also determine which AWS KMS key is used. To do this, choose the bucket in the table on the **S3 buckets** page. In the bucket details panel, under **Server-side encryption**, refer to the **KMS master key** field. This field shows the Amazon Resource Name (ARN) or unique identifier (key ID) for the key.

## Allowing Macie to use a customer managed AWS KMS key

If an S3 object is encrypted using a customer managed AWS KMS key (SSE-KMS encryption), Macie can decrypt and analyze the object only if Macie is allowed to use the KMS key. How to provide this access depends on whether the account that owns the key also owns the S3 bucket that stores the object:

- If the same account owns the KMS key and the bucket, a user of the account has to update the key's policy.
- If one account owns the KMS key and a different account owns the bucket, the account that owns the key has to allow cross-account access to the key.

This topic describes how to perform these tasks and provides examples for both scenarios.

### Allowing same-account access to a customer managed key

If the same account owns both the AWS KMS key and the bucket, a user of the account has to add a statement to the policy for the KMS key. The additional statement must allow the Macie service-linked role for the account to use the key to decrypt data. For detailed information about updating a key policy, see [Changing a key policy](#) in the *AWS Key Management Service Developer Guide*.

In the statement, the `Principal` element must specify the Amazon Resource Name (ARN) of the Macie service-linked role for the account that owns the KMS key and the bucket. If the account is in a manually enabled AWS Region, the ARN must include the appropriate Region code for the Region. For example, if the account is in the Middle East (Bahrain) Region, which has the Region code `me-south-1`, the `Principal` element must specify `arn:aws:iam::123456789012:role/aws-service-role/macie.me-south-1.amazonaws.com/AWSServiceRoleForAmazonMacie`, where `123456789012` is the account ID for the account.

The `Action` array must specify the `kms:Decrypt` action. This is the only AWS KMS action that Macie must be allowed to perform to decrypt an object that was encrypted with the key.

The following is an example of the statement to add to the policy for a KMS key.

```
{
  "Sid": "Allow the Macie service-linked role to use the key",
  "Effect": "Allow",
  "Principal": {
```

```
    "AWS": "arn:aws:iam::123456789012:role/aws-service-role/macie.amazonaws.com/  
AWSServiceRoleForAmazonMacie"  
  },  
  "Action": [  
    "kms:Decrypt"  
  ],  
  "Resource": "*"   
}
```

In the preceding example:

- The `AWS` field in the `Principal` element specifies the ARN of the Macie service-linked role (`AWSServiceRoleForAmazonMacie`) for the account. It allows the Macie service-linked role to perform the action specified by the policy statement. **123456789012** is an example account ID. Replace this ID with the account ID for the account that owns the KMS key and the bucket.
- The `Action` array specifies the action that the Macie service-linked role is allowed to perform using the KMS key—decrypt ciphertext that was encrypted with the key.

Where you add this statement to a key policy depends on the structure and elements that the policy currently contains. When you add the statement, ensure that the syntax is valid. Key policies use JSON format. This means that you have to also add a comma before or after the statement, depending on where you add the statement to the policy.

## Allowing cross-account access to a customer managed key

If one account owns the AWS KMS key (*key owner*) and a different account owns the bucket (*bucket owner*), the key owner has to provide the bucket owner with cross-account access to the KMS key. To do this, the key owner first ensures that the key's policy allows the bucket owner to both use the key and create a grant for the key. The bucket owner then creates a grant for the key. The grant delegates the relevant permissions to the Macie service-linked role for the bucket owner's account.

A *grant* is a policy instrument that allows AWS principals to use KMS keys in cryptographic operations if the conditions specified by the grant are met. To learn about grants, see [Using grants](#) in the *AWS Key Management Service Developer Guide*.

In the key policy, the key owner should ensure that the policy includes two statements. The first statement allows the bucket owner to use the key to decrypt data. The second statement allows the bucket owner to create a grant for the Macie service-linked role for the bucket owner's account. For detailed information about updating a key policy, see [Changing a key policy](#) in the *AWS Key Management Service Developer Guide*.

In the first statement, the `Principal` element must specify the ARN of the bucket owner's account. The `Action` array must specify the `kms:Decrypt` action. This is the only AWS KMS action that Macie must be allowed to perform to decrypt an object that was encrypted with the key.

The following is an example of this statement in the policy for a KMS key.

```
{  
  "Sid": "Allow account 111122223333 to use the key",  
  "Effect": "Allow",  
  "Principal": {  
    "AWS": "arn:aws:iam::111122223333:root"  
  },  
  "Action": [  
    "kms:Decrypt"  
  ],  
  "Resource": "*"   
}
```

In the preceding example:

- The `AWS` field in the `Principal` element specifies the ARN of the bucket owner's account (`111122223333`). It allows the bucket owner to perform the action specified by the policy statement. `111122223333` is an example account ID. Replace this ID with the account ID for the bucket owner's account.
- The `Action` array specifies the action that the bucket owner is allowed to perform using the KMS key—decrypt ciphertext that was encrypted with the key.

The second statement in the key policy allows the bucket owner to create a grant for the Macie service-linked role for their account. In this statement, the `Principal` element must specify the ARN of the bucket's owner's account. The `Action` array must specify the `kms:CreateGrant` action. A `Condition` element can filter access to the `kms:CreateGrant` action specified in the statement.

The following is an example of this statement in the policy for a KMS key.

```
{
  "Sid": "Allow account 111122223333 to create a grant",
  "Effect": "Allow",
  "Principal": {
    "AWS": "arn:aws:iam::111122223333:root"
  },
  "Action": [
    "kms:CreateGrant"
  ],
  "Resource": "*",
  "Condition": {
    "StringEquals": {
      "kms:GranteePrincipal": "arn:aws:iam::111122223333:role/aws-service-role/macie.amazonaws.com/AWSServiceRoleForAmazonMacie"
    }
  }
}
```

In the preceding example:

- The `AWS` field in the `Principal` element specifies the ARN of the bucket owner's account (`111122223333`). It allows the bucket owner to perform the action specified by the policy statement. `111122223333` is an example account ID. Replace this ID with the account ID for the bucket owner's account.
- The `Action` array specifies the action that the bucket owner is allowed to perform on the KMS key—create a grant for the key.
- The `Condition` element uses the `StringEquals` [condition operator](#) and the `kms:GranteePrincipal` [condition key](#) to filter access to the action specified by the policy statement. In this case, the bucket owner can create a grant only for the specified `GranteePrincipal`, which is the ARN of the Macie service-linked role for the bucket owner's account. In that ARN, `111122223333` is an example account ID. Replace this ID with the account ID for the bucket owner's account.

If the bucket owner's account is in a manually enabled AWS Region, also include the appropriate Region code in the ARN of the Macie service-linked role. For example, if the account is in the Middle East (Bahrain) Region, which has the Region code `me-south-1`, replace `macie.amazonaws.com` with `macie.me-south-1.amazonaws.com` in the ARN.

Where the key owner adds these statements to the key policy depends on the structure and elements that the policy currently contains. When the key owner adds the statement, they should ensure that the syntax is valid. Key policies use JSON format. This means that the key owner has to also add a comma before or after the statement, depending on where they add the statement to the policy.

After the key owner updates the key policy as necessary, the bucket owner must create a grant for the key. The grant delegates the relevant permissions to the Macie service-linked role for their (the bucket owner's) account. Before the bucket owner creates the grant, they should verify that they're allowed to perform the `kms:CreateGrant` action for their account. This action allows them to add a grant to an existing, customer managed KMS key.

To create the grant, the bucket owner can use the [CreateGrant](#) operation of the AWS Key Management Service API. When the bucket owner creates the grant, they should specify the following values for the required parameters:

- **GranteePrincipal** – The ARN of the Macie service-linked role (*AWSServiceRoleForAmazonMacie*) for their account. This value should be `arn:aws:iam::111122223333:role/aws-service-role/macie.amazonaws.com/AWSServiceRoleForAmazonMacie`, where **111122223333** is the account ID for the bucket owner's account.

If their account is in a manually enabled Region, the ARN must include the appropriate Region code. For example, if the account is in the Middle East (Bahrain) Region, which has the Region code *me-south-1*, the ARN should be `arn:aws:iam::111122223333:role/aws-service-role/macie.me-south-1.amazonaws.com/AWSServiceRoleForAmazonMacie`, where **111122223333** is the account ID for the bucket owner's account.

- **KeyId** – The ARN of the KMS key. For cross-account access to a KMS key, this value must be an ARN. It can't be a key ID.
- **Operations** – The AWS KMS decrypt action (`Decrypt`). This is the only AWS KMS action that Macie must be allowed to perform to decrypt an object that was encrypted with the KMS key.

The following example shows how to use the [AWS Command Line Interface \(AWS CLI\)](#) to create a grant for a customer managed KMS key. The example uses the `create-grant` command of the AWS Key Management Service API. The example is formatted for Microsoft Windows and it uses the caret (^) line-continuation character to improve readability.

```
C:\> aws kms create-grant ^
--key-id arn:aws:kms:us-east-1:123456789012:key/1234abcd-12ab-34cd-56ef-12345example ^
--grantee-principal arn:aws:iam::111122223333:role/aws-service-role/macie.amazonaws.com/
AWSServiceRoleForAmazonMacie ^
--operations "Decrypt"
```

Where:

- `key-id` specifies the ARN of the KMS key to apply the grant to.
- `grantee-principal` specifies the ARN of the Macie service-linked role for the account that's allowed to perform the operation specified by the grant. This value should match the ARN that's specified by the `kms:GranteePrincipal` condition of the second statement in the key policy.
- `operations` specifies the operation that the grant allows the specified principal to perform—decrypt ciphertext that was encrypted with the key.

If the command runs successfully, AWS KMS responds with output that's similar to the following.

```
{
  "GrantToken": "<grant token>",
  "GrantId": "1a2b3c4d2f5e69f440bae30eaec9570bb1fb7358824f9ddfa1aa5a0dab1a59b2"
}
```

Where `GrantToken` is a unique, non-secret, variable-length, base64-encoded string that represents the grant that was created, and `GrantId` is the unique identifier for the grant.

# Storing and retaining sensitive data discovery results with Amazon Macie

When Amazon Macie runs a sensitive data discovery job, it creates a record for each Amazon S3 object that you configure the job to analyze. This includes objects that don't contain sensitive data, and therefore don't produce a finding, and objects that Macie can't analyze due to issues such as permissions settings or use of an unsupported format. If an object does contain sensitive data, the record includes data from the corresponding finding. It provides additional information too, such as the location of as many as 1,000 occurrences of each type of sensitive data that Macie found in the object. Macie stores these records, referred to as *sensitive data discovery results*, for 90 days. To learn more about sensitive data discovery results, see [Reviewing job statistics and results \(p. 80\)](#).

To access your sensitive data discovery results and enable long-term storage and retention of them, configure Macie to store the results in an S3 bucket and encrypt them with an AWS Key Management Service (AWS KMS) key. If you do this, Macie writes your sensitive data discovery results to JSON Lines (.jsonl) files, which it adds to the S3 bucket as GNU Zip (.gz) files. The S3 bucket can then serve as a definitive, long-term repository for all of your sensitive data discovery results.

This topic walks you through the process of configuring this type of repository for your discovery results. The configuration is a combination of an S3 bucket that stores the results, an AWS KMS key that encrypts the results, and Macie settings that indicate which bucket and key to use.

When you configure the settings in Macie, your choices apply only to the current AWS Region. If your account is the Macie administrator account for an organization, your choices apply only to your account. They don't apply to any associated member accounts.

If you use Macie in multiple Regions, configure the repository settings for each Region in which you use Macie. If you prefer to store all discovery results for all Regions in one S3 bucket, you can do this by choosing the same bucket, located in one specific Region, for each (and every) Region in which you use Macie.

## Tasks

- [Step 1: Verify your permissions \(p. 96\)](#)
- [Step 2: Define the AWS KMS key and policy \(p. 97\)](#)
- [Step 3: Specify the S3 bucket to use \(p. 99\)](#)
- [Troubleshooting errors \(p. 103\)](#)

## Step 1: Verify your permissions

Before you configure a repository for your sensitive data discovery results, verify that you have the permissions that you need. You can do this by using the AWS Identity and Access Management (IAM) console:

1. Sign in to the AWS Management Console and open the IAM console at <https://console.aws.amazon.com/iam/>.
2. In the navigation pane, choose **Users**.
3. Choose your user name.

The **Permissions** tab lists all the IAM policies that are attached to your user name. Choose a policy to view its details. Then compare the information in the policy to the following list of actions that you must be allowed to perform in order to configure the repository.

## Macie

For Macie, verify that you're allowed to perform the following action:

```
macie2:PutClassificationExportConfiguration
```

This action allows you to add or change the repository settings in Macie.

## Amazon S3

For Amazon S3, verify that you're allowed to perform the following actions:

- `s3:CreateBucket`
- `s3:GetBucketLocation`
- `s3:ListAllMyBuckets`
- `s3:PutBucketAcl`
- `s3:PutBucketPolicy`
- `s3:PutBucketPublicAccessBlock`
- `s3:PutObject`

These actions allow you to access and configure an S3 bucket that can serve as the repository.

## AWS KMS

For AWS KMS, verify that you're allowed to perform the following action:

```
kms:ListAliases
```

This action allows you to retrieve information about AWS KMS keys that can encrypt the data in the repository. If you plan to create a new AWS KMS key to encrypt the data, you also need to be allowed to perform the following actions: `kms:CreateKey`, `kms:GetKeyPolicy`, and `kms:PutKeyPolicy`.

If you're not allowed to perform one or more of the preceding actions, ask your AWS administrator for assistance before you proceed to the next step.

# Step 2: Define the AWS KMS key and policy

When you configure Macie to store your sensitive data discovery results in an S3 bucket, you specify which AWS KMS key you want Macie to use to encrypt the results. The key must be a symmetric, customer managed AWS KMS key that's in the same AWS Region as the S3 bucket where you want to store the results. It can be an existing KMS key, or a new KMS key that you create before you configure the repository settings in Macie.

If you want to use a new KMS key, create the key before proceeding. To learn how, see [Creating keys](#) in the *AWS Key Management Service Developer Guide*. If you want to use an existing key that's owned by another account and you're allowed to use, obtain the Amazon Resource Name (ARN) of the key before proceeding. You'll need to enter this ARN when you configure the repository settings in Macie. To learn how to find a key's ARN, see [Finding the key ID and ARN](#) in the *AWS Key Management Service Developer Guide*.

After you determine which KMS key you want to use, give Macie permission to use the key. Otherwise, Macie won't be able to encrypt or store discovery results in the repository. To give Macie permission to use the key, change the key's policy.

## To change the key's policy

1. Open the AWS KMS console at <https://console.aws.amazon.com/kms>.
2. To change the AWS Region, use the Region selector in the upper-right corner of the page.

3. Choose the key that you want to use to encrypt the results.
4. On the **Key policy** tab, choose **Edit**.
5. Add the following statement to the policy:

```
{
  "Sid": "Allow Macie to use the key",
  "Effect": "Allow",
  "Principal": {
    "Service": "macie.amazonaws.com"
  },
  "Action": [
    "kms:GenerateDataKey",
    "kms:Encrypt"
  ],
  "Resource": "*"
}
```

#### Note

If you're using Macie in a manually enabled AWS Region, add the appropriate Region code to the value for the `Service` field in the statement. For example, if you're using Macie in the Middle East (Bahrain) Region, which has the Region code *me-south-1*, replace `macie.amazonaws.com` with `macie.me-south-1.amazonaws.com`.

When you add this statement to the policy, make sure that the syntax is valid. Policies use JSON format. This means that you need to also add a comma before or after the statement, depending on where you add the statement to the policy.

If you add the statement as the last statement, add a comma after the closing curly brace for the preceding section. If you add it as the first statement or between two existing statements, add a comma after the closing curly brace. The following examples show you how to add the statement to a default key policy.

The following example shows a default key policy that doesn't grant any additional permissions:

```
{
  "Id": "key-consolepolicy",
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Enable IAM user permissions",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::111122223333:root"
      },
      "Action": "kms:*",
      "Resource": "*"
    }
  ]
}
```

The following example shows you how to add the statement as the first statement in the policy:

```
{
  "Id": "key-consolepolicy",
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Allow Macie to use the key",
      "Effect": "Allow",
      "Principal": {
```

```
        "Service": "macie.amazonaws.com"
      },
      "Action": [
        "kms:GenerateDataKey",
        "kms:Encrypt"
      ],
      "Resource": "*"
    }, <-- Add a comma after this curly brace
    {
      "Sid": "Enable IAM user permissions",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::111122223333:root"
      },
      "Action": "kms:*",
      "Resource": "*"
    }
  ]
}
```

The following example shows you how to add the statement as the last statement in the policy:

```
{
  "Id": "key-consolepolicy",
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Enable IAM user permissions",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::111122223333:root"
      },
      "Action": "kms:*",
      "Resource": "*"
    }, <-- Add a comma after this curly brace
    {
      "Sid": "Allow Macie to use the key",
      "Effect": "Allow",
      "Principal": {
        "Service": "macie.amazonaws.com"
      },
      "Action": [
        "kms:GenerateDataKey",
        "kms:Encrypt"
      ],
      "Resource": "*"
    }
  ]
}
```

6. When you finish adding the statement, choose **Save changes**.

## Step 3: Specify the S3 bucket to use

After you verify your permissions and define the AWS KMS key to use, you're ready to specify which S3 bucket you want to use as the repository for your sensitive data discovery results. You have two options:

- **Use a new S3 bucket that Macie creates** – If you choose this option, Macie automatically creates a new S3 bucket for your discovery results. It also applies a bucket policy to the bucket. The policy allows Macie to create (put) objects in the bucket. To review this policy, choose **View policy** on the Amazon Macie console after you enter a name for the bucket.



- **Use an existing S3 bucket that you create** – If you prefer to store your discovery results in a particular S3 bucket that you create, create the bucket before you proceed. Then check the bucket's settings and update the bucket's policy to ensure that Macie can create (put) objects in the bucket. This topic explains which setting to check and how to update the policy. It also provides examples of the statements to add to the policy.

The following sections provide step-by-step instructions for each of these options. Choose the section for the option that you want.

## Use a new S3 bucket that Macie creates

If you prefer to use a new S3 bucket that Macie creates for you, the final step in the process is to configure the repository settings in Macie.

### To configure the repository settings in Macie

1. Open the Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, under **Settings**, choose **Discovery results**.
3. Under **Repository for sensitive data discovery results**, choose **Create bucket**.
4. In the **Create a bucket** box, enter a name for the bucket. The name must be unique across all S3 buckets. In addition, the name can consist only of lowercase letters, numbers, dots (.), and hyphens (-). For additional naming requirements, see [Bucket naming rules](#) in the *Amazon Simple Storage Service User Guide*.
5. (Optional) To specify a path prefix to use in the path to a location in the bucket, expand the **Advanced** section. Then, for **Data discovery result prefix**, enter the path prefix to use.  
  
When you enter a value, Macie updates the example below the field to show the path to the bucket location where it will store your discovery results.
6. For **Block all public access**, choose whether to enable all block public access settings for the bucket. For information about these settings, see [Blocking public access to your Amazon S3 storage](#) in the *Amazon Simple Storage Service User Guide*.
7. Under **KMS encryption**, specify the AWS KMS key that you want to use to encrypt the results:
  - To use a key for your own account, choose **Select a key from your account**. Then, in the **KMS key alias** list, choose the alias of the key to use.
  - To use a key that's owned by another account and you're allowed to use, choose **Enter the ARN of a key in another account**. Then, in the **KMS key ARN** field, enter the ARN of the key to use.

The key must be a symmetric, customer managed KMS key that's in the same Region as the S3 bucket.

8. When you finish entering the settings, choose **Save**. Macie then tests the settings to verify that they're correct. If any settings are incorrect, Macie displays an error message to help you address the issue.

After you save the repository settings, Macie adds existing discovery results for the preceding 90 days to the repository. Macie also starts adding new discovery results to the repository.

## Use an existing S3 bucket that you create

If you prefer to store your sensitive data discovery results in a particular S3 bucket that you create, create and configure the bucket before you configure the repository settings in Macie.

If you enabled Object Lock for the bucket, ensure that you disable the default retention setting for that feature. Otherwise, Macie won't be able to add your discovery results to the bucket. For information about this setting, see [Using S3 Object Lock](#) in the *Amazon Simple Storage Service User Guide*.

Then add a bucket policy that allows Macie to retrieve information about the bucket and create (put) objects in the bucket. You can then configure the repository settings in Macie.

### To add the bucket policy to the bucket

1. Open the Amazon S3 console at <https://console.aws.amazon.com/s3/>.
2. Choose the bucket that you want to store your discovery results in.
3. Choose the **Permissions** tab.
4. In the **Bucket policy** section, choose **Edit**.
5. Copy the following example policy to your clipboard:

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Allow Macie to use the GetBucketLocation operation",
      "Effect": "Allow",
      "Principal": {
        "Service": "macie.amazonaws.com"
      },
      "Action": "s3:GetBucketLocation",
      "Resource": "arn:aws:s3:::myBucketName"
    },
    {
      "Sid": "Allow Macie to upload objects to the bucket",
      "Effect": "Allow",
      "Principal": {
        "Service": "macie.amazonaws.com"
      },
      "Action": "s3:PutObject",
      "Resource": "arn:aws:s3:::myBucketName/[optional prefix]/*"
    },
    {
      "Sid": "Deny unencrypted object uploads. This is optional",
      "Effect": "Deny",
      "Principal": {
        "Service": "macie.amazonaws.com"
      },
      "Action": "s3:PutObject",
      "Resource": "arn:aws:s3:::myBucketName/[optional prefix]/*",
      "Condition": {
        "StringNotEquals": {
          "s3:x-amz-server-side-encryption": "aws:kms"
        }
      }
    },
    {
      "Sid": "Deny incorrect encryption headers. This is optional",
      "Effect": "Deny",
      "Principal": {
        "Service": "macie.amazonaws.com"
      },
      "Action": "s3:PutObject",
      "Resource": "arn:aws:s3:::myBucketName/[optional prefix]/*",
      "Condition": {
        "StringNotEquals": {
          "s3:x-amz-server-side-encryption-aws-kms-key-id":
            "arn:aws:kms:Region:111122223333:key/KMSKeyId"
        }
      }
    },
    {
      "Sid": "Deny non-HTTPS access",
```

```
    "Effect": "Deny",
    "Principal": "*",
    "Action": "s3:*",
    "Resource": "arn:aws:s3:::myBucketName/*",
    "Condition": {
      "Bool": {
        "aws:SecureTransport": "false"
      }
    }
  }
]
```

6. Paste the example policy in the **Bucket policy** editor on the Amazon S3 console. Then replace the placeholder values with the correct values for your environment, where:

- **myBucketName** is the name of the bucket.
- **Region** is the AWS Region that hosts the AWS KMS key to use for encryption of the discovery results.
- **111122223333** is the account ID for your AWS account, or the AWS account that owns the KMS key to use for encryption of the discovery results.
- **KMSKeyId** is the key ID of the KMS key to use for encryption of the discovery results.

If you're using Macie in a manually enabled AWS Region, also add the appropriate Region code to the value for the **Service** field in each statement that specifies the Macie service principal. For example, if you're using Macie in the Middle East (Bahrain) Region, which has the Region code *me-south-1*, replace `macie.amazonaws.com` with `macie.me-south-1.amazonaws.com` in each applicable statement.

7. When you finish updating the bucket policy, choose **Save changes**.

### Important

If you change the bucket path after you configure the repository settings in Macie, you have to update the bucket policy. Otherwise, Macie won't be allowed to add discovery results to the bucket.

You can now configure the repository settings in Macie.

### To configure the repository settings in Macie

1. Open the Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, under **Settings**, choose **Discovery results**.
3. Under **Repository for sensitive data discovery results**, choose **Existing bucket**.
4. For **Choose a bucket**, select the bucket that you want to store your discovery results in.
5. (Optional) To specify a path prefix to use in the path to a location in the bucket, expand the **Advanced** section. Then, for **Data discovery result prefix**, enter the path prefix to use.

When you enter a value, Macie updates the example below the field to show the path to the bucket location where it will store your discovery results.

6. Under **KMS encryption**, specify the AWS KMS key that you want to use to encrypt the results:
  - To use a key for your own account, choose **Select a key from your account**. Then, in the **KMS key alias** list, choose the alias of the key to use.
  - To use a key that's owned by another account and you're allowed to use, choose **Enter the ARN of a key in another account**. Then, in the **KMS key ARN** field, enter the ARN of the key to use.

The key must be a symmetric, customer managed KMS key in the same Region as the S3 bucket that you specified.

7. When you finish entering the settings, choose **Save**. Macie then tests the settings to verify that they're correct. If any settings are incorrect, Macie displays an error message to help you address the issue.

After you save the repository settings, Macie adds existing discovery results for the preceding 90 days to the repository. Macie also starts adding new discovery results to the repository.

## Troubleshooting errors

If an error occurs when Macie tries to add sensitive data discovery results to the repository, Macie displays an error message on the **Repository for sensitive data discovery results** page of the console. In addition, we notify you by sending email to the address that's associated with your AWS account. If you don't address the error, Macie stores backups of your discovery results for up to 90 days.

Errors typically occur because Macie loses access to the repository—for example, the S3 bucket is deleted or the permissions settings for the bucket are changed. They also occur if the AWS KMS key that's used to encrypt the results becomes inaccessible. If an error occurs, use the information in this topic as a guide to walk through possible causes and solutions for the error. For example, review the policy for the KMS key and confirm that it's still correct.

After you address the error, update the configuration settings in Macie. Macie then starts adding new discovery results to the repository. Macie also adds any existing results that it created and stored while the error existed (for up to 90 days).

# Analyzing Amazon Macie findings

Amazon Macie generates a finding each time it detects a potential policy violation for an Amazon Simple Storage Service (Amazon S3) bucket or it discovers sensitive data in an S3 object. A *finding* is a detailed report of a potential policy violation or sensitive data that Macie found. Each finding provides a severity rating, information about the affected resource, and additional details, such as when and how Macie found the issue. Macie stores your findings for 90 days.

You can view, analyze, and manage findings in the following ways.

## Amazon Macie console

The **Findings** pages on the Amazon Macie console list your findings and provide detailed information for individual findings. These pages also provide options for grouping, filtering, and sorting findings, and for creating and managing [suppression rules \(p. 152\)](#). Suppression rules can help you streamline your analysis of findings.

## Amazon Macie API

You can also use the Amazon Macie API to query and retrieve findings data. You can query the data by using the Amazon Macie REST API, the AWS Command Line Interface (AWS CLI), or another AWS SDK or tool of your choice. To query the data, you send a request to the Amazon Macie API and use supported parameters to specify which findings you want to retrieve. After you submit your query, Macie returns the results in a JSON response. You can then pass the results to another service or application for deeper analysis, long-term storage, or reporting. For more information, see the [Amazon Macie API Reference](#).

## Amazon EventBridge

To further support integration with other services and systems, such as monitoring or event management systems, Macie publishes findings as events to Amazon EventBridge. EventBridge, formerly called Amazon CloudWatch Events, is a serverless event bus service that can deliver a stream of real-time data from your own applications, software as a service (SaaS) applications, and AWS services such as Macie. It can route that data to targets such as AWS Lambda functions, Amazon Simple Notification Service topics, and Amazon Kinesis streams. To learn about this service, see the [Amazon EventBridge User Guide](#).

Macie automatically publishes events to EventBridge for new findings. It also publishes events automatically for subsequent occurrences of existing policy findings. Because the notifications are structured as EventBridge events, you can more easily monitor, analyze, and act upon findings by using other services and tools. For example, you might use EventBridge to automatically send specific types of new findings to an AWS Lambda function that, in turn, processes and sends the data to your security incident and event management (SIEM) system. In addition to automated processing, use of EventBridge events helps ensure longer-term retention of findings data. To learn about using EventBridge events for findings, see [Monitoring and processing findings \(p. 162\)](#).

## AWS Security Hub

For additional, broader analysis of your organization's security posture, you can also review and analyze findings by using AWS Security Hub. Security Hub is a service that provides you with a comprehensive view of your security state across your AWS environment and helps you check your environment against security industry standards and best practices. To learn about this service, see the [AWS Security Hub User Guide](#). To learn about how Macie publishes findings to Security Hub, see [Monitoring and processing findings \(p. 162\)](#).

In addition to findings, Macie creates sensitive data discovery results for S3 objects that you configure it to analyze as part of a sensitive data discovery job. A *sensitive data discovery result* is a record that logs details about a job's analysis of an object. This includes objects that don't contain sensitive data, and

therefore don't produce a finding, and objects that Macie can't analyze due to issues such as permission settings for a bucket. You can't access these results directly on the Amazon Macie console or through the Amazon Macie API. Instead, you configure Macie to store the results in an S3 bucket. You can then access the results in that bucket. For more information, see [Storing and retaining sensitive data discovery results](#) (p. 96).

#### Topics

- [Types of Amazon Macie findings](#) (p. 105)
- [Viewing findings on the Amazon Macie console](#) (p. 107)
- [Locating sensitive data with Amazon Macie findings](#) (p. 108)
- [Filtering Amazon Macie findings](#) (p. 117)
- [Suppressing Amazon Macie findings](#) (p. 152)
- [Severity scoring for Amazon Macie findings](#) (p. 157)

## Types of Amazon Macie findings

Amazon Macie generates two categories of findings: *policy findings* and *sensitive data findings*. A *policy finding* is a detailed report of a potential policy violation for an Amazon Simple Storage Service (Amazon S3) bucket. Macie generates these findings as part of its ongoing monitoring activities for your Amazon S3 data. A *sensitive data finding* is a detailed report of sensitive data in an S3 object. Macie generates these findings when it discovers sensitive data in S3 objects that you configure it to analyze as part of a sensitive data discovery job.

#### Topics

- [Policy findings](#) (p. 105)
- [Sensitive data findings](#) (p. 106)

## Policy findings

Macie generates policy findings when the policies or settings for an S3 bucket are changed in a way that reduces the security of the bucket and its objects. Macie does this only if the change occurs after you enable Macie for your Amazon Web Services account.

For example, if default encryption is disabled for a bucket after you enable Macie, Macie generates a **Policy:IAMUser/S3BucketEncryptionDisabled** finding for the bucket. However, if default encryption was disabled for a bucket when you enabled Macie and default encryption continues to be disabled, Macie doesn't generate a **Policy:IAMUser/S3BucketEncryptionDisabled** finding for the bucket.

Macie can generate the following types of policy findings for an S3 bucket.

#### **Policy:IAMUser/S3BlockPublicAccessDisabled**

Block public access settings were disabled for the bucket. Access to the bucket is controlled only by access control lists (ACLs) and bucket policies.

To learn about block public access settings for S3 buckets, see [Blocking public access to your Amazon S3 storage](#) in the *Amazon Simple Storage Service User Guide*.

#### **Policy:IAMUser/S3BucketEncryptionDisabled**

Default encryption was disabled for the bucket. By default, Amazon S3 won't automatically encrypt new objects when they're added to the bucket.

To learn about default encryption settings for S3 buckets, see [Setting default server-side encryption behavior for Amazon S3 buckets](#) in the *Amazon Simple Storage Service User Guide*.

**Policy:IAMUser/S3BucketPublic**

An ACL or bucket policy for the bucket was changed to allow access by anonymous users or by all authenticated AWS Identity and Access Management (IAM) users or roles.

To learn about ACLs and bucket policies for S3 buckets, see [Identity and access management in Amazon S3](#) in the *Amazon Simple Storage Service User Guide*.

**Policy:IAMUser/S3BucketReplicatedExternally**

Data replication was enabled and configured to replicate objects from the bucket to an Amazon Web Services account that isn't part of your organization. An *organization* is a set of Macie accounts that are centrally managed as a group of related accounts through AWS Organizations or by Macie invitation.

To learn about replication settings for S3 buckets, see [Replicating objects](#) in the *Amazon Simple Storage Service User Guide*.

**Policy:IAMUser/S3BucketSharedExternally**

An ACL or bucket policy for the bucket was changed to allow the bucket to be shared with an Amazon Web Services account that isn't part of your organization. An *organization* is a set of Macie accounts that are centrally managed as a group of related accounts through AWS Organizations or by Macie invitation.

To learn about ACLs and bucket policies for S3 buckets, see [Identity and access management in Amazon S3](#) in the *Amazon Simple Storage Service User Guide*.

## Sensitive data findings

Macie generates sensitive data findings when it discovers sensitive data in S3 objects that you configure it to analyze as part of a sensitive data discovery job. Macie can generate the following types of sensitive data findings for an object.

**SensitiveData:S3Object/Credentials**

The object contains credentials, such as private keys or AWS secret keys.

**SensitiveData:S3Object/CustomIdentifier**

The object contains content that matches one or more custom data identifiers. The object might include more than one type of sensitive data.

**SensitiveData:S3Object/Financial**

The object contains financial information, such as credit card numbers or bank account numbers.

**SensitiveData:S3Object/Multiple**

The object contains more than one category of sensitive data—any combination of credentials, financial information, personal information, or content that matches one or more custom data identifiers.

**SensitiveData:S3Object/Personal**

The object contains personally identifiable information (such as full names or mailing addresses), personal health information (such as health insurance or medical identification numbers), or a combination of the two.

For detailed information about the types of sensitive data that Macie can detect, see [Using managed data identifiers \(p. 37\)](#). For information about the types of S3 objects that Macie can analyze, see [Supported file and storage formats \(p. 89\)](#).

# Viewing findings on the Amazon Macie console

Amazon Macie monitors your AWS environment and generates policy findings when it detects potential policy violations for your Amazon Simple Storage Service (Amazon S3) buckets. Macie generates sensitive data findings when it discovers sensitive data in S3 objects that you configure it to analyze as part of a sensitive data discovery job. Macie stores your policy and sensitive data findings for 90 days.

By using the Amazon Macie console, you can review and analyze findings, and view the details of individual findings. Each finding provides a severity rating, information about the affected resource, and additional details, such as the exact nature of the issue, and when and how Macie found the issue.

To help you streamline your analysis, the console offers several options for building custom views of findings.

## Use predefined groupings

Use specific pages to view findings that are grouped by criteria such as affected S3 bucket, finding type, or sensitive data discovery job. With these pages, you can view aggregated statistics for each group, such as the count of findings by severity. You can also drill down to view the details of individual findings in a group, and you can apply filters to refine your analysis.

For example, if you view all findings grouped by S3 bucket and see that a particular S3 bucket has a policy violation, you can quickly determine whether the bucket also contains sensitive data. To do this, choose **By bucket** in the navigation pane (under **Findings**), and then choose the bucket. In the details panel that appears, the **Findings by type** section lists the types of findings that apply to the bucket. To investigate a specific type, choose the number for the type. Macie displays a table of all the findings that both match the selected type and apply to the bucket. To refine the results, filter the table.

## Apply attribute-based filters

Use specific finding attributes to include or exclude certain findings from a **Findings** table. A *finding attribute* is a field that stores specific data for a finding, such as finding type, severity, or the name of the S3 bucket that the finding applies to. If you filter a table, you can more easily identify findings that have specific characteristics. Then you can drill down to view the details of those findings.

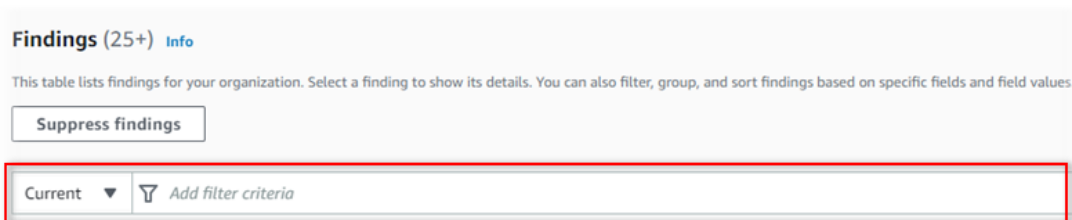
For example, to review all of your policy findings, add filter criteria for the **Category** field. To refine your view and include only a specific type of policy finding, add filter criteria for the **Finding type** field. To then review the details of a particular finding, choose the finding. The details panel displays information for the finding.

You can also sort findings in ascending or descending order by certain fields. To do this, click the column heading for the field. To change the sort order, click the column heading again.

## To view findings on the console

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Findings**. The **Findings** page displays current findings for your account in the current AWS Region. By default, this doesn't include findings that were suppressed by a [suppression rule](#) (p. 152).
3. (Optional) To view and pivot on findings by a predefined logical group, choose **By bucket**, **By type**, or **By job** in the navigation pane (under **Findings**), and then choose an item in the table. In the details panel, choose the link for the field to pivot on.
4. (Optional) To filter the findings by specific criteria, use the filter bar above the table:







- To display findings that were suppressed by a suppression rule, choose **Current** in the filter bar. Then choose **Archived** to display only suppressed findings, or choose **All** to display both current and suppressed findings.
- To display only those findings that have a specific attribute, place your cursor in the filter bar and add a filter condition for the attribute. To further refine the results, add conditions for additional attributes. For information about using filter conditions, see [Creating and applying filters to findings \(p. 123\)](#).
- To remove a filter condition, choose the remove condition icon (⊗) in the filter box.

To save your filter settings, choose **Save rule** in the filter bar. Then enter a name and, optionally, a description for the settings. When you finish, choose **Save**.

5. (Optional) To sort the findings by a specific field, click the column heading for the field. To change the sort order, click the column heading again.
6. To view the details of a specific finding, choose any field other than the check box for the finding. The details panel displays information for the finding.

#### Tip

You can use the details panel to pivot and drill down on certain fields by choosing a magnifying glass for the field. Choose  to show findings with the same value, or choose  to show findings with other values.

For a sensitive data finding, you can also use the details panel to locate occurrences of sensitive data in the affected object, or navigate to the corresponding sensitive data discovery result for the finding:

- To locate occurrences of sensitive data, choose a link in an **Occurrences** field. Macie displays information (in JSON format) about where Macie found the data. To learn more, see [Locating sensitive data with findings \(p. 108\)](#).
- To navigate to the corresponding sensitive data discovery result, choose the link in the **Detailed result location** field. Macie opens the Amazon S3 console and displays the file or folder that contains the discovery result. To learn more, see [Reviewing job statistics and results \(p. 80\)](#).

You can also download and save the details of one or more findings as a JSON file. To do this, select the check box for each finding that you want to download and save. Then choose **Export (JSON)** from the **Actions** menu at the top of the **Findings** page. In the window that appears, choose **Download**.

## Locating sensitive data with Amazon Macie findings

When you run a sensitive data discovery job, Amazon Macie captures details about the location of each occurrence of sensitive data that it finds in an Amazon S3 object. This includes sensitive data that Macie

detects using [managed data identifiers](#) (p. 37), and data that matches any [custom data identifiers](#) (p. 53) that you configure a sensitive data discovery job to use.

With sensitive data findings, you can view these details for as many as 15 occurrences of sensitive data that Macie detects when it runs a job. The details provide insight into the breadth of the categories and types of sensitive data that specific S3 buckets and objects contain. They can also help you locate individual occurrences of sensitive data and determine whether to perform a deeper investigation of specific buckets and objects.

To help you locate an occurrence of sensitive data, a finding can provide details such as:

- The column and row number for a cell or field in a Microsoft Excel workbook, CSV file, or TSV file.
- The path to a field or array in a JSON or JSON Lines file.
- The line number for a line in a non-binary text file other than a CSV, JSON, JSON Lines, or TSV file—for example, an HTML, TXT, or XML file.
- The page number for a page in an Adobe Portable Document Format (PDF) file.
- The record index and the path to a field in a record in an Apache Avro object container or Apache Parquet file.

You can access these details by using the Amazon Macie console and the Amazon Macie API. You can also access these details in findings that Macie publishes to other AWS services, both Amazon EventBridge and AWS Security Hub.

If an S3 object contains many occurrences of sensitive data, you can also use a finding to navigate to the corresponding sensitive data discovery result for the finding. Unlike a sensitive data finding, a sensitive data discovery result provides detailed location data for as many as 1,000 occurrences of each type of sensitive data that Macie detects in an object. If an S3 object is an archive file, such as a .tar or .zip file, a sensitive data discovery result also provides detailed location data for occurrences of sensitive data in individual files that Macie extracts from the archive file. (Macie doesn't include this information in sensitive data findings.) For more information about sensitive data discovery results, see [Reviewing job statistics and results](#) (p. 80). Macie uses the same schema for location data in sensitive data findings and sensitive data discovery results.

The topics in this section explain how to locate occurrences of sensitive data by using sensitive data findings and the Amazon Macie console. They also explain the schema that Macie uses to store and report the location of individual occurrences of sensitive data. To access location data programmatically, you can use the [Findings Descriptions](#) resource of the Amazon Macie API. To learn how to access the data in findings that Macie publishes to other AWS services, see [Monitoring and processing findings](#) (p. 162).

#### Topics

- [Locating occurrences of sensitive data](#) (p. 109)
- [JSON schema for sensitive data locations](#) (p. 110)
- [JSON details and examples for sensitive data locations](#) (p. 112)

## Locating occurrences of sensitive data

When you run a sensitive data discovery job, Macie performs a deep inspection of the latest version of each S3 object that you configure the job to analyze. Macie also uses a *depth-first search* algorithm to populate the job's findings with details about the location of 1–15 occurrences of the sensitive data that Macie detects. These occurrences provide insight into the categories and types of sensitive data that the affected S3 buckets and objects contain. You can use these details to locate individual occurrences of sensitive data and determine whether to perform a deeper investigation of specific buckets and objects.

### To locate occurrences of sensitive data

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Findings**.

#### Tip

You can also use the **Jobs** page to display all the findings from a particular job. To do this, choose **Jobs** in the navigation pane, and then choose the name of the job. At the top of the details panel, choose **Show results**, and then choose **Show findings**.

3. On the **Findings** page, choose the finding for the sensitive data that you want to locate. The details panel displays information for the finding.
4. In the details panel, scroll to the **Details** section. This section provides information about the categories and types of sensitive data that Macie found in the affected S3 object.

If the finding includes details about where Macie found a type of sensitive data, an **Occurrences** field appears and summarizes those details, as shown in the following image.

Financial information	
Credit card number	29
Occurrences of credit card number	5 line ranges
Personal information	
Address	28
Occurrences of address	5 line ranges
Name	32
Occurrences of name	5 line ranges

To show the details for a specific type of sensitive data, choose the link in the **Occurrences** field. Macie opens a new window and displays the details in JSON format. To then save the details as a JSON file, choose **Download** and specify a name and location for the file.

5. (Optional) To save all the finding's details as a JSON file, choose the finding's identifier (**Finding ID**) at the top of the details panel. Macie opens a new window and displays all the details in JSON format. Choose **Download**, and then specify a name and location for the file.

To access details about the location of as many as 1,000 occurrences of each type of sensitive data in an affected object, you can refer to the corresponding sensitive data discovery result for the finding. To help you do this, the details panel provides a link to the discovery result. In the details panel, scroll to the **Details** section of the panel, and then choose the link in the **Detailed result location** field. Macie opens the Amazon S3 console and displays the file or folder that contains the discovery result. To learn more about these results, see [Reviewing job statistics and results](#) (p. 80).

## JSON schema for sensitive data locations

Macie uses standardized JSON structures to store information about where it finds sensitive data in S3 objects. These structures are used by sensitive data findings and sensitive data discovery results. For sensitive data findings, the structures are part of the JSON schema for Macie findings. To view the complete JSON schema for Macie findings, see [Findings Descriptions](#) in the *Amazon Macie API Reference*.

The JSON schema for a sensitive data finding includes one `customDataIdentifiers` object and one `sensitiveData` object. The `customDataIdentifiers` object provides details about data that Macie detected using [custom data identifiers](#) (p. 53). The `sensitiveData` object provides details about sensitive data that Macie detected using [managed data identifiers](#) (p. 37).

Each `customDataIdentifiers` and `sensitiveData` object contains one or more `detections` arrays:

- In a `customDataIdentifiers` object, the `detections` array indicates the custom data identifiers that detected the data and produced the finding. For each custom data identifier, the array also indicates the number of occurrences of the data that the identifier detected. It can also indicate the location of the data that the identifier detected.
- In a `sensitiveData` object, a `detections` array indicates the types of sensitive data that Macie detected using managed data identifiers. For each type of sensitive data, the array also indicates the number of occurrences of the data, and it can indicate the location of the data.

For a sensitive data finding, a `detections` array can include 1–15 `occurrences` objects. Each `occurrences` object specifies where Macie found individual occurrences of a specific type of sensitive data.

For example, the following `detections` array indicates the location of three occurrences of sensitive data (US Social Security numbers) in a CSV file.

```
"sensitiveData": [
  {
    "category": "PERSONAL_INFORMATION",
    "detections": [
      {
        "count": 30,
        "occurrences": {
          "cells": [
            {
              "cellReference": null,
              "column": 1,
              "columnName": "SSN",
              "row": 2
            },
            {
              "cellReference": null,
              "column": 1,
              "columnName": "SSN",
              "row": 3
            },
            {
              "cellReference": null,
              "column": 1,
              "columnName": "SSN",
              "row": 4
            }
          ]
        }
      },
      {
        "type": "USA_SOCIAL_SECURITY_NUMBER"
      }
    ]
  }
]
```

The location and number of `occurrences` objects in a `detections` array varies based on the categories, types, and number of occurrences of sensitive data that Macie detects when it runs a sensitive data discovery job. This variation occurs because Macie includes location data for only 1–15 occurrences of the sensitive data that it detects when it runs a job. These 1–15 occurrences are indicative of the categories and types of sensitive data that the affected S3 buckets and objects contain.

An `occurrences` object can contain any the following structures, depending on an S3 object's file type or storage format:

- **cells** array – This array applies to Microsoft Excel workbooks, CSV files, and TSV files. An object in this array specifies a cell or field that contains an occurrence of sensitive data.
- **lineRanges** array – This array applies to non-binary text files other than CSV, JSON, JSON Lines, and TSV files—for example, HTML, TXT, and XML files. An object in this array specifies a line or an inclusive range of lines that contains an occurrence of sensitive data, and the position of the data on the specified line or lines.

In certain cases, an object in a **lineRanges** array specifies the location of sensitive data in a file type or storage format that's supported by another type of array. Those cases are: sensitive data in an unstructured section of an otherwise structured file, such as a comment in a file; sensitive data in a malformed file that Macie analyzes as plaintext; and, a CSV or TSV file that has one or more column names that contain sensitive data.

- **offsetRanges** array – This array is reserved for future use. If this array is present, the value for it is always null.
- **pages** array – This array applies to Adobe Portable Document Format (PDF) files. An object in this array specifies a page that contains an occurrence of sensitive data.
- **records** array – This array applies to Apache Avro object containers, Apache Parquet files, JSON files, and JSON Lines files. For Avro object containers and Parquet files, an object in this array specifies a record index and the path to a field in a record that contains an occurrence of sensitive data. For JSON and JSON Lines files, an object in this array specifies the path to a field or array that contains an occurrence of sensitive data. For JSON Lines files, it also specifies the index of the line that contains the data.

The contents of these arrays vary based on an affected S3 object's file type or storage format and its contents. The next topic provides details and examples of each array.

## JSON details and examples for sensitive data locations

Macie tailors the contents of the JSON structures that it uses to indicate the location of sensitive data in specific types of files and content. The following topics explain and provide examples of these structures.

### Topics

- [Cells array \(p. 112\)](#)
- [LineRanges array \(p. 113\)](#)
- [Pages array \(p. 115\)](#)
- [Records array \(p. 115\)](#)

For a complete list of JSON structures that can be included in a sensitive data finding, see [Findings Descriptions](#) in the *Amazon Macie API Reference*.

### Cells array

**Applies to:** Microsoft Excel workbooks, CSV files, and TSV files

In a **cells** array, a **Cell** object specifies a cell or field that contains an occurrence of sensitive data. The following table describes the purpose of each field in a **Cell** object.

Field	Type	Description
<code>cellReference</code>	String	The location of the cell, as an absolute cell reference, that

Field	Type	Description
		contains the sensitive data. This field applies only to Excel workbooks. This value is null for CSV and TSV files.
column	Integer	The column number of the column that contains the sensitive data. For an Excel workbook, this value correlates to the alphabetical character(s) for a column identifier—for example, 1 for column A, 2 for column B, and so on.
columnName	String	The name of the column that contains the sensitive data, if available.
row	Integer	The row number of the row that contains the sensitive data.

The following example shows the structure of a `Cell` object that reports an occurrence of sensitive data in a CSV file.

```
"cells": [
  {
    "cellReference": null,
    "column": 3,
    "columnName": "SSN",
    "row": 5
  }
]
```

In the preceding example, the finding indicates that the field in the fifth row of the third column (named *SSN*) of the file contains sensitive data.

The following example shows the structure of a `Cell` object that reports an occurrence of sensitive data in an Excel workbook.

```
"cells": [
  {
    "cellReference": "Sheet2!C5",
    "column": 3,
    "columnName": "SSN",
    "row": 5
  }
]
```

In the preceding example, the finding indicates that the worksheet named *Sheet2* in the workbook contains sensitive data. In that worksheet, the sensitive data is in the cell in the fifth row of the third column (column C, named *SSN*).

## LineRanges array

**Applies to:** Non-binary text files other than CSV, JSON, JSON Lines, and TSV files—for example, HTML, TXT, and XML files

In a `lineRanges` array, a `Range` object specifies a line or an inclusive range of lines that contains an occurrence of sensitive data, and the position of the data on the specified line or lines.

This object is often empty for file types that are supported by other types of arrays in `occurrences` objects. Exceptions are:

- Data in unstructured sections of an otherwise structured file, such as a comment in a file.
- Data in a malformed file that Macie analyzes as plaintext.
- A CSV or TSV file that has one or more column names that contain sensitive data.

The following table describes the purpose of each field in a `Range` object of a `lineRanges` array.

Field	Type	Description
<code>end</code>	Integer	The number of lines from the beginning of the file to the end of the sensitive data.
<code>start</code>	Integer	The number of lines from the beginning of the file to the beginning of the sensitive data.
<code>startColumn</code>	Integer	The number of characters, with spaces and starting from 1, from the beginning of the first line that contains the sensitive data ( <code>start</code> ) to the beginning of the sensitive data.

The following example shows the structure of a `Range` object that reports an occurrence of sensitive data that's stored on a single line in a TXT file.

```
"lineRanges": [  
  {  
    "end": 1,  
    "start": 1,  
    "startColumn": 119  
  }  
]
```

In the preceding example, the finding indicates that the first line of the file contains a complete occurrence of sensitive data (a mailing address). The first character in the occurrence is 119 characters (with spaces) from the beginning of that line.

The following example shows the structure of a `Range` object that reports an occurrence of sensitive data that spans multiple lines in a TXT file.

```
"lineRanges": [  
  {  
    "end": 54,  
    "start": 51,  
    "startColumn": 1  
  }  
]
```

In the preceding example, the finding indicates that lines 51 through 54 of the file contain an occurrence of sensitive data (a mailing address). The first character in the occurrence is the first character on line 51 of the file.

## Pages array

**Applies to:** Adobe Portable Document Format (PDF) files

In a `pages` array, a `Page` object specifies a page that contains an occurrence of sensitive data. The object contains a `pageNumber` field. The `pageNumber` field stores an integer that specifies the page number of the page that contains the sensitive data.

The following example shows the structure of a `Page` object that reports an occurrence of sensitive data in a PDF file.

```
"pages": [  
  {  
    "pageNumber": 10  
  }  
]
```

In the preceding example, the finding indicates that page 10 of the file contains sensitive data.

## Records array

**Applies to:** Apache Avro object containers, Apache Parquet files, JSON files, and JSON Lines files

For an Avro object container or a Parquet file, a `Record` object in a `records` array specifies a record index and the path to a field in a record that contains an occurrence of sensitive data. For JSON and JSON Lines files, a `Record` object specifies the path to a field or array that contains an occurrence of sensitive data. For JSON Lines files, it also specifies the index of the line that contains the data.

The following table describes the purpose of each field in a `Record` object.

Field	Type	Description
<code>jsonPath</code>	String	<p>The path, as a JSONPath expression, to the sensitive data.</p> <p>For an Avro object container or a Parquet file, this is the path to the field in the record (<code>recordIndex</code>) that contains the data. For a JSON or JSON Lines file, this is the path to the field or array that contains the data. If the data is a value in an array, the path also indicates which value contains the data.</p> <p>If Macie detects sensitive data in the name of any element in the path, Macie omits the <code>jsonPath</code> field from a <code>Record</code> object. If the name of a path element exceeds 20 characters, Macie truncates the name by removing</p>



Field	Type	Description
		characters from the beginning of the name. If the resulting full path exceeds 250 characters, Macie also truncates the path, starting with the first element in the path, until the path contains 250 or fewer characters.
recordIndex	Integer	For an Avro object container or a Parquet file, the record index, starting from 0, for the record that contains the sensitive data. For a JSON Lines file, the line index, starting from 0, for the line that contains the sensitive data. This value is always 0 for JSON files.

The following example shows the structure of a `Record` object that reports an occurrence of sensitive data in a Parquet file. In this example, Macie truncated the name of the field that contains the data, specified in the `jsonPath` field, to meet the character limit.

```
"records": [
  {
    "jsonPath": "$['...hijklmnopqrstuvwxyz']",
    "recordIndex": 7663
  }
]
```

In the preceding example, the finding indicates that the record of index 7663 (record number 7664) contains sensitive data. In that record, the sensitive data is in the field whose name ends with `hijklmnopqrstuvwxyz`. The full JSON path to the field in the record is `$.abcdefghijklmnopqrstuvwxyz`.

The following example also shows the structure of a `Record` object that reports an occurrence of sensitive data in a Parquet file. In this example, Macie truncated both the full path and the name of the field that contains the data.

```
"records": [
  {
    "jsonPath":
"$..usssn2.usssn3.usssn4.usssn5.usssn6.usssnfield7.usssn8.usssn9.usssn10.usssn11.usssn12.usssn13.usssn14.usssn15.usssn16.usssn17.usssn18.usssn19.usssn20.usssn21.usssn22.usssn23.usssn24.usssn25.usssn26.usssn27.usssn28.usssn29['abcdefghijklmnopqrstuvwxyz']",
    "recordIndex": 2335
  }
]
```

In the preceding example, the finding indicates that the record of index 2335 (record number 2336) contains sensitive data. In that record, the sensitive data is in the field whose name ends with `hijklmnopqrstuvwxyz`. The full JSON path to the field in the record is: `$['1234567890'].usssn1.usssn2.usssn3.usssn4.usssn5.usssn6.usssnfield7.usssn8.usssn9.usssn10.usssn11.usssn12.usssn13.usssn14.usssn15.usssn16.usssn17.usssn18.usssn19.usssn20.usssn21.usssn22.usssn23.usssn24.usssn25.usssn26.usssn27.usssn28.usssn29['abcdefghijklmnopqrstuvwxyz']`

The following example shows the structure of a `Record` object that reports an occurrence of sensitive data in a JSON file. In this example, the sensitive data is a specific value in an array.

```
"records": [  
  {  
    "jsonPath": "$.access.key[2]",  
    "recordIndex": 0  
  }  
]
```

In the preceding example, the finding indicates that the second value in an array named `key` contains sensitive data. The array is a child of an object named `access`.

The following example shows the structure of a `Record` object that reports an occurrence of sensitive data in a JSON Lines file.

```
"records": [  
  {  
    "jsonPath": "$.access.key",  
    "recordIndex": 3  
  }  
]
```

In the preceding example, the finding indicates that the third value (line) in the file contains sensitive data. In that line, the sensitive data is in a field named `key`, which is a child of an object named `access`.

## Filtering Amazon Macie findings

To perform targeted analysis and to analyze findings more efficiently, you can filter Amazon Macie findings. With filters, you build custom views and queries for findings, which can help you identify and focus on findings that have specific characteristics. Use the Amazon Macie console to filter findings, or submit queries programmatically using the Amazon Macie API.

When you create a filter, you use specific attributes of findings to define criteria for including or excluding findings from a view or from query results. A *finding attribute* is a field that stores specific data for a finding, such as severity, type, or the name of the S3 bucket that a finding applies to.

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

- An attribute-based field, such as **Severity** or **Finding type**.
- An operator, such as *equals* or *not equals*.
- One or more values. The type and number of values depends on the field and operator that you choose.

If you create a filter that you want to use again, you can save it as a *filter rule*. A *filter rule* is a set of filter criteria that you create and save to reapply when you view findings on the Amazon Macie console.

You can also save a filter as a *suppression rule*. A *suppression rule* is a set of filter criteria that you create and save to automatically archive findings that meet the criteria for a rule. To learn about suppression rules, see [Suppressing findings \(p. 152\)](#).

### Topics

- [Fundamentals of filtering findings \(p. 118\)](#)
- [Creating and applying filters to findings \(p. 123\)](#)
- [Creating and managing filter rules for findings \(p. 129\)](#)
- [Fields for filtering findings \(p. 134\)](#)

## Fundamentals of filtering findings

When you create a filter, keep the following features and guidelines in mind. Also note that filtered results are limited to current findings in the current AWS Region. Amazon Macie stores your findings for 90 days in each AWS Region.

### Topics

- [Using multiple conditions in a filter \(p. 118\)](#)
- [Specifying values for fields \(p. 118\)](#)
- [Specifying multiple values for a field \(p. 120\)](#)
- [Using operators in conditions \(p. 120\)](#)

## Using multiple conditions in a filter

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

- An attribute-based field, such as **Severity** or **Finding type**. For a list of fields that you can use, see [Fields for filtering findings \(p. 134\)](#).
- An operator, such as *equals* or *not equals*. For a list of operators that you can use, see [Using operators in conditions \(p. 120\)](#).
- One or more values. The type and number of values depends on the field and operator that you choose.

If a filter contains multiple conditions, Macie uses AND logic to join the conditions and evaluate the filter criteria. This means that a finding meets the filter criteria only if it matches *all* the conditions in the filter.

For example, if you add a condition to include only high-severity findings and add another condition to include only sensitive data findings, Macie returns all high-severity, sensitive data findings. In other words, Macie excludes all policy findings and all medium-severity and low-severity sensitive data findings.

You can use a field only once in a filter. However, you can specify multiple values for many fields.

For example, if a condition uses the **Severity** field to include only high-severity findings, you can't use the **Severity** field in another condition to include medium-severity or low-severity findings. Instead, specify multiple values for the existing condition, or use a different operator for the existing condition. For example, to include all medium-severity and high-severity findings, add a **Severity equals Medium, High** condition or add a **Severity not equals Low** condition.

## Specifying values for fields

When you specify a value for a field, the value has to conform to the underlying data type for the field. Depending on the field, you can specify one of the following types of values.

### Array of text (strings)

Specifies a list of text (string) values for a field. Each string correlates to a predefined or existing value for a field—for example, *High* for the **Severity** field, *SensitiveData:S3Object/Financial* for the **Finding type** field, or the name of an S3 bucket for the **S3 bucket name** field.

If you use an array, note the following:

- Values are case sensitive.

- You can't specify partial values or use wildcard characters in values. You have to specify a complete, valid value for the field.

For example, to filter findings for an S3 bucket named *my-S3-bucket*, enter **my-S3-bucket** as the value for the **S3 bucket name** field. If you enter any other value, such as **my-s3-bucket** or **my-S3**, Macie won't return findings for the bucket.

For a list of valid values for each field, see [Fields for filtering findings \(p. 134\)](#).

You can specify as many as 50 values in an array. How you specify the values depends on whether you use the Amazon Macie console or the Amazon Macie API, as discussed in [Specifying multiple values for a field \(p. 120\)](#).

### Boolean

Specifies one of two mutually exclusive values for a field.

If you use the Amazon Macie console to specify this type of value, the console provides a list of values to choose from. If you use the Amazon Macie API, specify `true` or `false` for the value.

### Date/Time (and time ranges)

Specifies an absolute date and time for a field. If you specify this type of value, you have to specify both a date and time.

On the Amazon Macie console, date and time values are in your local time zone and use 24-hour notation. In all other contexts, these values are in Coordinated Universal Time (UTC) and extended ISO 8601 format—for example `2020-09-01T14:31:13Z` for 2:31:13 PM UTC September 1, 2020.

If a field stores a date/time value, you can use the field to define a fixed or relative time range. For example, you can include only those findings that were created between two specific dates and times, or only those findings that were created before or after a specific date and time. How you define a time range depends on whether you use the Amazon Macie console or the Amazon Macie API:

- On the console, use a date picker or enter text directly in the **From** and **To** boxes.
- With the API, define a fixed time range by adding a condition that specifies the first date and time in the range, and add another condition that specifies the last date and time in the range. If you do this, Macie uses AND logic to join the conditions. To define a relative time range, add one condition that specifies the first or last date and time in the range. Specify the values as Unix timestamps in milliseconds—for example, `1604616572653` for 22:49:32 UTC November 5, 2020.

On the console, time ranges are inclusive. With the API, time ranges can be inclusive or exclusive, depending on the operator that you choose.

### Number (and numeric ranges)

Specifies a long integer for a field.

If a field stores a numeric value, you can use the field to define a fixed or relative numeric range. For example, you can include only those findings that report 50-90 occurrences of sensitive data in an S3 object. How you define a numeric range depends on whether you use the Amazon Macie console or the Amazon Macie API:

- On the console, use the **From** and **To** boxes to enter the lowest and highest numbers in the range, respectively.
- With the API, define a fixed numeric range by adding a condition that specifies the lowest number in the range, and add another condition that specifies the highest number in the range. If you do this, Macie uses AND logic to join the conditions. To define a relative numeric range, add one condition that specifies the lowest or highest number in the range.

On the console, numeric ranges are inclusive. With the API, numeric ranges can be inclusive or exclusive, depending on the operator that you choose.

### Text (string)

Specifies a single text (string) value for a field. The string correlates to a predefined or existing value for a field—for example, *High* for the **Severity** field, the name of an S3 bucket for the **S3 bucket name** field, or the unique identifier for a sensitive data discovery job for the **Job ID** field.

If you specify a single text string, note the following:

- Values are case sensitive.
- You can't use partial values or use wildcard characters in values. You have to specify a complete, valid value for the field.

For example, to filter findings for an S3 bucket named *my-S3-bucket*, enter **my-S3-bucket** as the value for the **S3 bucket name** field. If you enter any other value, such as **my-s3-bucket** or **my-S3**, Macie won't return findings for the bucket.

For a list of valid values for each field, see [Fields for filtering findings \(p. 134\)](#).

## Specifying multiple values for a field

With certain fields and operators, you can specify multiple values for a field. If you do this, Macie uses OR logic to join the values and evaluate the filter criteria. This means that a finding meets the criteria if it has *any* of the values for the field.

For example, if you add a condition to include findings where the value for the **Finding type** field equals *SensitiveData:S3Object/Financial*, *SensitiveData:S3Object/Personal*, Macie returns sensitive data findings for S3 objects that contain only financial data, and S3 objects that contain only personal information. In other words, Macie excludes all policy findings. Macie also excludes all sensitive data findings for objects that contain other types of sensitive data or multiple types of sensitive data.

The exception is conditions that use the *eqExactMatch* operator. For this operator, Macie uses AND logic to join the values and evaluate the filter criteria. This means that a finding meets the criteria only if it has *all* the values for the field and *only* those values for the field. To learn more about this operator, see [Using operators in conditions \(p. 120\)](#).

How you specify multiple values for a field depends on whether you use the Amazon Macie API or the Amazon Macie console. With the API, you use an array that lists the values.

On the console, you typically choose the values from a list. However, for some fields, you have to add a distinct condition for each value. For example, to include findings for data that Macie detected using certain custom data identifiers, do the following:

1. Place your cursor in the filter bar, choose the **Customer data identifier detection name** field, enter the name of a custom data identifier, and then choose **Apply**.
2. Repeat the preceding step for each additional custom data identifier that you want to specify for the filter.

For a list of fields that you need to do this for, see [Fields for filtering findings \(p. 134\)](#).

## Using operators in conditions

You can use the following types of operators in individual conditions.

### Equals (eq)

Matches (=) any value specified for the field. You can use the *equals* operator with the following types of values: array of text (strings), Boolean, date/time, number, and text (string).

For many fields, you can use this operator and specify as many as 50 values for the field. If you do this, Macie uses OR logic to join the values. This means that a finding meets the criteria if it has *any* of the values specified for the field.

For example:

- To include findings that report occurrences of financial information, personal information, or both financial and personal information, add a condition that uses the **Sensitive data category** field and this operator, and specify *Financial information* and *Personal information* as the values for the field.
- To include findings that report occurrences of credit card numbers, mailing addresses, or both credit card numbers and mailing addresses, add a condition for the **Sensitive data detection type** field, use this operator, and specify *CREDIT\_CARD\_NUMBER* and *ADDRESS* as the values for the field.

If you use the Amazon Macie API to define a condition that uses this operator with a date/time value, specify the value as a Unix timestamp in milliseconds—for example, 1604616572653 for 22:49:32 UTC November 5, 2020.

#### **Equals exact match (eqExactMatch)**

Exclusively matches all the values specified for the field. You can use the *equals exact match* operator with a select set of fields.

If you use this operator and specify multiple values for a field, Macie uses AND logic to join the values. This means that a finding meets the criteria only if it has *all* the values specified for the field and *only* those values for the field. You can specify as many as 50 values for the field.

For example:

- To include findings that report occurrences of credit card numbers and no other type of sensitive data, add a condition for the **Sensitive data detection type** field, use this operator, and specify *CREDIT\_CARD\_NUMBER* as the only value for the field.
- To include findings that report occurrences of both credit card numbers and mailing addresses (and no other types of sensitive data), add a condition for the **Sensitive data detection type** field, use this operator, and specify *CREDIT\_CARD\_NUMBER* and *ADDRESS* as the values for the field.

Because Macie uses AND logic to join the values for a field, you can't use this operator in combination with any other operators for the same field. In other words, if you use the *equals exact match* operator with a field in one condition, you have to use it in all other conditions that use the same field.

Like other operators, you can use the *equals exact match* operator in more than one condition in a filter. If you do this, Macie uses AND logic to join the conditions and evaluate the filter. This means that a finding meets the filter criteria only if it has *all* the values specified by *all* the conditions in the filter.

For example, to include findings that were created after a certain time, report occurrences of credit card numbers, and don't report any other type of sensitive data, do the following:

1. Add a condition that uses the **Created at** field, uses the *greater than* operator, and specifies the starting date and time for the filter.
2. Add another condition that uses the **Sensitive data detection type** field, uses the *equals exact match* operator, and specifies *CREDIT\_CARD\_NUMBER* as the only value for the field.

You can use the *equals exact match* operator with the following fields:

- Customer data identifier detection ARN (`customDataIdentifiers.detections.arn`)
- Customer data identifier detection name (`customDataIdentifiers.detections.name`)
- S3 bucket tag key (`resourcesAffected.s3Bucket.tags.key`)
- S3 bucket tag value (`resourcesAffected.s3Bucket.tags.value`)
- S3 object tag key (`resourcesAffected.s3Object.tags.key`)

- S3 object tag value (`resourcesAffected.s3Object.tags.value`)
- Sensitive data detection type (`sensitiveData.detections.type`)
- Sensitive data category (`sensitiveData.category`)

In the preceding list, the parenthetical name uses dot notation to indicate the name of the field in JSON representations of findings and the Amazon Macie API.

#### Greater than (gt)

Is greater than ( $>$ ) the value specified for the field. You can use the *greater than* operator with number and date/time values.

For example, to include only those findings that report more than 90 occurrences of sensitive data in an S3 object, add a condition that uses the **Sensitive data total count** field and this operator, and specify 90 as the value for the field. To do this on the Amazon Macie console, enter **91** in the **From** box, don't enter a value in the **To** box, and then choose **Apply**. Numeric and time-based comparisons are inclusive on the console.

If you use the Amazon Macie API to define a time range that uses this operator, you have to specify the date/time values as Unix timestamps in milliseconds—for example, 1604616572653 for 22:49:32 UTC November 5, 2020.

#### Greater than or equal to (gte)

Is greater than or equal to ( $\geq$ ) the value specified for the field. You can use the *greater than or equal to* operator with number and date/time values.

For example, to include only those findings that report 90 or more occurrences of sensitive data in an S3 object, add a condition that uses the **Sensitive data total count** field and this operator, and specify 90 as the value for the field. To do this on the Amazon Macie console, enter **90** in the **From** box, don't enter a value in the **To** box, and then choose **Apply**.

If you use the Amazon Macie API to define a time range that uses this operator, you have to specify the date/time values as Unix timestamps in milliseconds—for example, 1604616572653 for 22:49:32 UTC November 5, 2020.

#### Less than (lt)

Is less than ( $<$ ) the value specified for the field. You can use the *less than* operator with number and date/time values.

For example, to include only those findings that report fewer than 90 occurrences of sensitive data in an S3 object, add a condition that uses the **Sensitive data total count** field and this operator, and specify 90 as the value for the field. To do this on the Amazon Macie console, enter **89** in the **To** box, don't enter a value in the **From** box, and then choose **Apply**. Numeric and time-based comparisons are inclusive on the console.

If you use the Amazon Macie API to define a time range that uses this operator, you have to specify the date/time values as Unix timestamps in milliseconds—for example, 1604616572653 for 22:49:32 UTC November 5, 2020.

#### Less than or equal to (lte)

Is less than or equal to ( $\leq$ ) the value specified for the field. You can use the *less than or equal to* operator with number and date/time values.

For example, to include only those findings that report 90 or fewer occurrences of sensitive data in an S3 object, add a condition that uses the **Sensitive data total count** field and this operator, and specify 90 as the value for the field. To do this on the Amazon Macie console, enter **90** in the **To** box, don't enter a value in the **From** box, and then choose **Apply**.

If you use the Amazon Macie API to define a time range that uses this operator, you have to specify the date/time values as Unix timestamps in milliseconds—for example, 1604616572653 for 22:49:32 UTC November 5, 2020.

### Not equals (neq)

Doesn't match ( $\neq$ ) any value specified for the field. You can use the *not equals* operator with the following types of values: array of text (strings), Boolean, date/time, number, and text (string).

For many fields, you can use this operator and specify as many as 50 values for the field. If you do this, Macie uses OR logic to join the values. This means that a finding meets the criteria if it doesn't have *any* of the values specified for the field.

For example:

- To exclude findings that report occurrences of financial information, personal information, or both financial and personal information, add a condition that uses the **Sensitive data category** field and this operator, and specify *Financial information* and *Personal information* as the values for the field.
- To exclude findings that report occurrences of credit card numbers, add a condition for the **Sensitive data detection type** field, use this operator, and specify *CREDIT\_CARD\_NUMBER* as the value for the field.
- To exclude findings that report occurrences of credit card numbers, mailing addresses, or both credit card numbers and mailing addresses, add a condition for the **Sensitive data detection type** field, use this operator, and specify *CREDIT\_CARD\_NUMBER* and *ADDRESS* as the values for the field.

If you use the Amazon Macie API to define a condition that uses this operator with a date/time value, specify the value as a Unix timestamp in milliseconds—for example, 1604616572653 for 22:49:32 UTC November 5, 2020.

## Creating and applying filters to findings

To identify and focus on findings that have specific characteristics, you can filter findings on the Amazon Macie console and in queries that you submit programmatically using the Amazon Macie API. When you create a filter, you use specific attributes of findings to define criteria for including or excluding findings from a view or from query results. A *finding attribute* is a field that stores specific data for a finding, such as severity, type, or the name of the S3 bucket that a finding applies to.

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

- An attribute-based field, such as **Severity** or **Finding type**.
- An operator, such as *equals* or *not equals*.
- One or more values. The type and number of values depends on the field and operator that you choose.

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

### Topics

- [Filtering findings on the Amazon Macie console \(p. 123\)](#)
- [Filtering findings programmatically with the Amazon Macie API \(p. 126\)](#)

## Filtering findings on the Amazon Macie console

If you use the Amazon Macie console to filter findings, Macie provides options to help you choose fields, operators, and values for individual conditions. You access these options by using the filter bar on **Findings** pages, as shown in the following image.



### Findings (25+) [Info](#)

This table lists findings for your organization. Select a finding to show its details. You can also filter, group, and sort findings based on specific fields and field values.

[Suppress findings](#)

Current ▼  [Add filter criteria](#)

When you place your cursor in the filter bar, Macie displays a list of fields that you can use in filter conditions. The fields are organized by logical category. For example, the **Common fields** category includes fields that apply to any type of finding, and the **Classification fields** category includes fields that apply only to sensitive data findings. The fields are sorted alphabetically within each category.

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

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

After you add a condition, Macie applies the criteria for the condition and adds the condition to a filter box in the filter bar, as shown in the following image.

Current ▼  ● **Severity: Medium, High**  [Add filter](#)

In this example, the condition is configured to include all medium-severity and high-severity findings, and to exclude all low-severity findings. It returns findings where the value for the **Severity** field *equals* **Medium** or **High**.

#### Tip

For many fields, you can change a condition's operator from *equals* to *not equals* by choosing the equals icon (●) in a filter box. If you do this, Macie changes the operator to *not equals* and displays the not equals icon (⊘) in the filter box. To switch to the *equals* operator again, choose the not equals icon.

As you add more conditions, Macie applies their criteria and adds them to the filter bar. You can refer to the filter bar at any time to see which criteria you've applied. To remove a condition, choose the remove condition icon (⊘) in the filter box for the condition.

### To filter findings using the console

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Findings**.
3. (Optional) To first view and pivot on findings by a predefined logical group, choose **By bucket**, **By type**, or **By job** in the navigation pane (under **Findings**), and then choose an item in the table. In the details panel, choose the link for the field to pivot on.
4. (Optional) To display findings that were suppressed by a [suppression rule \(p. 152\)](#), choose **Current** in the filter bar. Then choose **Archived** to display only suppressed findings, or choose **All** to display both current and suppressed findings.
5. To add a filter condition:

- a. Place your cursor in the filter bar, and then choose the field to use for the condition. For information about the fields that you can use, see [Fields for filtering findings \(p. 134\)](#).
- b. Enter the appropriate type of value for the field. For detailed information about the different types of values, see [Specifying values for fields \(p. 118\)](#).

#### Array of text (strings)

For this type of value, Macie often provides a list of values to choose from. If this is the case, select each value that you want to use in the condition.

If Macie doesn't provide a list of values, enter a complete, valid value for the field. To specify additional values for the field, choose **Apply**, and then add another condition for each additional value.

Note that values are case sensitive. In addition, you can't use partial values or wildcard characters in values. For example, to filter findings for an S3 bucket named *my-S3-bucket*, enter **my-S3-bucket** as the value for the **S3 bucket name** field. If you enter any other value, such as **my-s3-bucket** or **my-S3**, Macie won't return findings for the bucket.

#### Boolean

For this type of value, Macie provides a list of values to choose from. Select the value that you want to use in the condition.

#### Date/Time (time ranges)

For this type of value, use the **From** and **To** boxes to define an inclusive time range:

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

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

#### Number (numeric ranges)

For this type of value, use the **From** and **To** boxes to enter one or more integers that define an inclusive, fixed or relative numeric range.

#### Text (string) values

For this type of value, enter a complete, valid value for the field.

Note that values are case sensitive. In addition, you can't use partial values or wildcard characters in values. For example, to filter findings for an S3 bucket named *my-S3-bucket*, enter **my-S3-bucket** as the value for the **S3 bucket name** field. If you enter any other value, such as **my-s3-bucket** or **my-S3**, Macie won't return findings for the bucket.

- c. When you finish adding values for the field, choose **Apply**. Macie applies the filter criteria and adds the condition to a filter box in the filter bar.
6. Repeat step 5 for each additional condition that you want to add.
7. To remove a condition, choose the remove condition icon (✕) in the filter box for the condition.
8. To change a condition, remove the condition by choosing the remove condition icon (✕) in the filter box for the condition. Then repeat step 5 to add a condition with the correct settings.

If you want to subsequently use this set of conditions again, you can save the filter as a filter rule. To do this, choose **Save rule** in the filter bar. Then enter a name and, optionally, a description for the rule. When you finish, choose **Save**.

## Filtering findings programmatically with the Amazon Macie API

To filter findings programmatically, specify filter criteria in queries that you submit using the [ListFindings](#) or [GetFindingStatistics](#) operation of the Amazon Macie API. The **ListFindings** operation returns an array of finding IDs, one ID for each finding that meets the filter criteria. The **GetFindingStatistics** operation returns aggregated statistical data about all the findings that meet the filter criteria, grouped by a field that you specify in your request.

Note that the **ListFindings** and **GetFindingStatistics** operations are different from operations that you use to [suppress findings](#) (p. 152). Unlike suppression operations, which also specify filter criteria, the **ListFindings** and **GetFindingStatistics** operations only query findings data. They don't perform any action on findings that meet filter criteria. To suppress findings, use the [Findings Filters](#) resource of the Amazon Macie API.

To specify filter criteria in a query, include a map of filter conditions in your request. For each condition, specify a field, an operator, and one or more values for the field. The type and number of values depends on the field and operator that you choose. For information about the fields, operators, and types of values that you can use in a condition, see [Fields for filtering findings](#) (p. 134), [Using operators in conditions](#) (p. 120), and [Specifying values for fields](#) (p. 118).

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

### Examples

- [Example 1: Filter findings based on severity](#) (p. 127)
- [Example 2: Filter findings based on sensitive data category](#) (p. 127)
- [Example 3: Filter findings based on a fixed time range](#) (p. 127)
- [Example 4: Filter findings based on suppression status](#) (p. 128)
- [Example 5: Filter findings based on multiple fields and types of values](#) (p. 128)

The examples use the `list-findings` command. If an example runs successfully, Macie returns a `findingIds` array. The array lists the unique identifier for each finding that meets the filter criteria, as shown in the following example.

```
{
  "findingIds": [
    "1f1c2d74db5d8caa76859ec52example",
    "6cfa9ac820dd6d55cad30d851example",
    "702a6fd8750e567d1a3a63138example",
    "826e94e2a820312f9f964cf60example",
    "274511c3fdcd87010a19a3a42example"
  ]
}
```

If no findings meet the filter criteria, Macie returns an empty `findingIds` array.

```
{
  "findingIds": []
}
```

## Example 1: Filter findings based on severity

This example uses the `list-findings` command to retrieve finding IDs for all of your high-severity and medium-severity findings in the current AWS Region.

For Linux, macOS, or Unix:

```
$ aws macie2 list-findings --finding-criteria '{"criterion":{"severity.description":{"eq":["High","Medium"]}}}'
```

For Microsoft Windows:

```
C:\> aws macie2 list-findings --finding-criteria={"criterion":{"severity.description":{"eq":["High","Medium"]}}}
```

Where:

- `severity.description` specifies the JSON name of the **Severity** field.
- `eq` specifies the *equals* operator.
- `High` and `Medium` are an array of enumerated values for the **Severity** field.

## Example 2: Filter findings based on sensitive data category

This example uses the `list-findings` command to retrieve finding IDs for all of your sensitive data findings that are in the current Region and report occurrences of financial data (and no other categories of sensitive data) in S3 objects.

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

```
$ aws macie2 list-findings \  
--finding-criteria '{"criterion":{"classificationDetails.result.sensitiveData.category":{"eqExactMatch":["FINANCIAL_INFORMATION"]}}}'
```

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

```
C:\> aws macie2 list-findings ^  
--finding-criteria={"criterion":{"classificationDetails.result.sensitiveData.category\  
{"eqExactMatch":["FINANCIAL_INFORMATION"]}}}
```

Where:

- `classificationDetails.result.sensitiveData.category` specifies the JSON name of the **Sensitive data category** field.
- `eqExactMatch` specifies the *equals exact match* operator.
- `FINANCIAL_INFORMATION` is an enumerated value for the **Sensitive data category** field.

## Example 3: Filter findings based on a fixed time range

This example uses the `list-findings` command to retrieve finding IDs for all of your findings that are in the current Region and were created between 07:00 UTC October 5, 2020, and 07:00 UTC November 5, 2020 (inclusively).

For Linux, macOS, or Unix:

```
$ aws macie2 list-findings --finding-criteria '{"criterion":{"createdAt":{"gte":"1601881200000","lte":"1604559600000"}}}'
```

For Microsoft Windows:

```
C:\> aws macie2 list-findings --finding-criteria={"criterion":{"createdAt":{"gte":"1601881200000","lte":"1604559600000"}}}
```

Where:

- `createdAt` specifies the JSON name of the **Created at** field.
- `gte` specifies the *greater than or equal to* operator.
- `1601881200000` is the first date and time (as a Unix timestamp in milliseconds) in the time range.
- `lte` specifies the *less than or equal to* operator.
- `1604559600000` is the last date and time (as a Unix timestamp in milliseconds) in the time range.

### Example 4: Filter findings based on suppression status

This example uses the `list-findings` command to retrieve finding IDs for all of your findings that are in the current Region and were suppressed (automatically archived) by a suppression rule.

For Linux, macOS, or Unix:

```
$ aws macie2 list-findings --finding-criteria '{"criterion":{"archived":{"eq":["true"]}}}'
```

For Microsoft Windows:

```
C:\> aws macie2 list-findings --finding-criteria={"criterion":{"archived":{"eq":["true"]}}}
```

Where:

- `archived` specifies the JSON name of the **Archived** field.
- `eq` specifies the *equals* operator.
- `true` is a Boolean value for the **Archived** field.

### Example 5: Filter findings based on multiple fields and types of values

This example uses the `list-findings` command to retrieve finding IDs for all of your sensitive data findings that are in the current Region and meet the following criteria: were created between 07:00 UTC October 5, 2020, and 07:00 UTC November 5, 2020 (exclusively); report occurrences of financial data and no other categories of sensitive data in S3 objects; and weren't suppressed (automatically archived) by a suppression rule.

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

```
$ aws macie2 list-findings \  
--finding-criteria '{"criterion":{"createdAt":{"gt":"1601881200000","lt":"1604559600000"},"classificationDetails.result.sensitiveData.category":{"eqExactMatch":["FINANCIAL_INFORMATION"]},"archived":{"eq":["false"]}}}'
```

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

```
C:\> aws macie2 list-findings ^
--finding-criteria={"criterion":{"createdAt":{"gt":1601881200000,
"lt":1604559600000},"classificationDetails.result.sensitiveData.category":
{"eqExactMatch":["FINANCIAL_INFORMATION"]},"archived":{"eq":["false"]}}}
```

Where:

- `createdAt` specifies the JSON name of the **Created at** field, and:
  - `gt` specifies the *greater than or equal to* operator.
  - `1601881200000` is the first date and time (as a Unix timestamp in milliseconds) in the time range.
  - `lt` specifies the *less than or equal to* operator.
  - `1604559600000` is the last date and time (as a Unix timestamp in milliseconds) in the time range.
- `classificationDetails.result.sensitiveData.category` specifies the JSON name of the **Sensitive data category** field, and:
  - `eqExactMatch` specifies the *equals exact match* operator.
  - `FINANCIAL_INFORMATION` is an enumerated value for the field.
- `archived` specifies the JSON name of the **Archived** field, and:
  - `eq` specifies the *equals* operator.
  - `false` is a Boolean value for the field.

## Creating and managing filter rules for findings

A *filter rule* is a set of filter criteria that you create and save to use again when you view findings on the Amazon Macie console. Filter rules can help you perform consistent analysis of findings that have specific characteristics. For example, you might create one filter rule for analyzing all high-severity policy findings for S3 buckets that contain unencrypted objects, and another filter rule for analyzing all high-severity sensitive data findings that report specific types of sensitive data.

Note that filter rules are different from suppression rules. A *suppression rule* is a set of filter criteria that you create and save to automatically archive findings that meet the criteria of a rule. Although both types of rules store and apply filter criteria, a filter rule doesn't perform any action on findings that meet the criteria of a rule. Instead, a filter rule only determines which findings appear on the console after you apply the rule. For information about suppression rules, see [Suppressing findings \(p. 152\)](#).

To create and manage filter rules, you can use the Amazon Macie console or the Amazon Macie API. The following topics explain how. For the API, the topics explain how to perform these tasks with the [AWS Command Line Interface \(AWS CLI\)](#). You can also perform these tasks by sending HTTPS requests directly to Macie, or by using a current version of another AWS command line tool or an AWS SDK. For information about AWS tools and SDKs, see [Tools to Build on AWS](#).

### Topics

- [Creating filter rules \(p. 129\)](#)
- [Applying filter rules \(p. 131\)](#)
- [Changing filter rules \(p. 132\)](#)
- [Deleting filter rules \(p. 133\)](#)

## Creating filter rules

When you create a filter rule, you specify filter criteria, a name, and, optionally, a description for the rule. You can create a filter rule using the Amazon Macie console or the Amazon Macie API.

## Console

Follow these steps to create a filter rule by using the Amazon Macie console.

### To create a filter rule

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Findings**.

#### Tip

To use an existing filter rule as a starting point, choose the rule from the **Saved rules** list.

You can also streamline creation of a rule by first pivoting and drilling down on findings by a predefined logical group. If you do this, Macie automatically creates and applies the appropriate filter conditions, which can be a helpful starting point for creating a rule. To do this, choose **By bucket**, **By type**, or **By job** in the navigation pane (under **Findings**), and then choose an item in the table. In the details panel, choose the link for the field to pivot on.

3. In the filter bar, add conditions that define the filter criteria for the rule. To learn how, see [Creating and applying filters to findings \(p. 123\)](#).
4. When you finish defining filter criteria for the rule, choose **Save rule** in the filter bar.



5. Under **Filter rule**, enter a name and, optionally, a description for the rule.
6. Choose **Save**.

## AWS CLI

To create a filter rule by using the AWS CLI, run the `create-findings-filter` command and specify the appropriate values for the required parameters. For the `action` parameter, specify `NOOP` to ensure that Macie doesn't suppress (automatically archive) findings that meet the criteria of the rule.

For the `criterion` parameter, specify a map of conditions that define the filter criteria for the rule. In the map, each condition should specify a field, an operator, and one or more values for the field. The type and number of values depends on the field and operator that you choose. For information about the fields, operators, and types of values that you can use in a condition, see [Fields for filtering findings \(p. 134\)](#), [Using operators in conditions \(p. 120\)](#), and [Specifying values for fields \(p. 118\)](#).

The following examples create a filter rule that returns all sensitive data findings that are in the current AWS Region and report occurrences of personal information (and no other categories of sensitive data) in S3 objects.

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

```
$ aws macie2 create-findings-filter \
--action NOOP \
--name my_filter_rule \
--finding-criteria '{"criterion":
{"classificationDetails.result.sensitiveData.category":{"eqExactMatch":
["PERSONAL_INFORMATION"]}}}'
```

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

```
C:\> aws macie2 create-findings-filter ^
--action NOOP ^
--name my_filter_rule ^
--finding-criteria={"criterion":
{"classificationDetails.result.sensitiveData.category":{"eqExactMatch":
{"PERSONAL_INFORMATION"}}}}
```

Where:

- `my_filter_rule` is the custom name for the rule.
- `criterion` is a map of filter conditions for the rule:
  - `classificationDetails.result.sensitiveData.category` is the JSON name of the **Sensitive data category** field.
  - `eqExactMatch` specifies the *equals exact match* operator.
  - `PERSONAL_INFORMATION` is an enumerated value for the **Sensitive data category** field.

If the command runs successfully, you receive output similar to the following.

```
{
  "arn": "arn:aws:macie2:us-west-2:123456789012:findings-filter/9b2b4508-aa2f-4940-
b347-d1451example",
  "id": "9b2b4508-aa2f-4940-b347-d1451example"
}
```

Where `arn` is the Amazon Resource Name (ARN) of the filter rule that was created, and `id` is the unique identifier for the rule.

For additional examples of filter criteria, see [Filtering findings programmatically with the Amazon Macie API \(p. 126\)](#).

## Applying filter rules

When you apply a filter rule, Macie uses the criteria of the rule to determine which findings to include or exclude from your view of findings on the console. Macie also displays the criteria in the filter bar.

Note that filter rules are designed for use with the Amazon Macie console. You can't use them directly in queries that you submit programmatically using the Amazon Macie API. However, if you're using the API to query findings, you can retrieve the filter criteria for a rule by using the [GetFindingsFilter](#) operation of the API, and then add the criteria to your query. For information about specifying filter criteria in a query, see [Creating and applying filters to findings \(p. 123\)](#).

Follow these steps to filter findings on the console by applying a filter rule.

### To apply a filter rule

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Findings**.
3. In the **Saved rules** list, choose the filter rule that you want to apply. Macie applies the criteria of the rule and displays the criteria in the filter bar.
4. (Optional) To refine the criteria, use the filter bar to add or remove filter conditions. If you do this, your changes won't affect the settings for the rule. Macie won't save any of your changes unless you explicitly save them as a new rule.
5. To apply a different filter rule, repeat step 3.



After you apply a filter rule, you can quickly remove all of its filter criteria from your view by choosing the **X** in the filter bar.


## Changing filter rules

You can change the settings for a filter rule at any time using the Amazon Macie console or the Amazon Macie API.

### Console

Follow these steps to change the settings for an existing filter rule by using the Amazon Macie console.

#### To change a filter rule

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Findings**.
3. In the **Saved rules** list, choose the edit icon () next to the filter rule that you want to change.
4. Do any of the following:
  - To change the name of the rule, enter a new name in the **Name** box under **Filter rule**.
  - To change the description of the rule, enter a new description in the **Description** box under **Filter rule**.
  - To change the filter criteria of the rule, use the filter bar to enter conditions for the criteria that you want. To learn how, see [Creating and applying filters to findings \(p. 123\)](#).
5. When you finish making changes, choose **Save**.

### AWS CLI

To change a filter rule by using the AWS CLI, run the `update-findings-filter` command and use the supported parameters to specify a new value for each setting that you want to change. For the `id` parameter, specify the unique identifier for the rule to change. You can get this identifier by running the `list-findings-filters` command to retrieve a list of filter and suppression rules for your account.

The following example changes the name of an existing filter rule.

```
C:\> aws macie2 update-findings-filter --id 9b2b4508-aa2f-4940-b347-d1451example --  
name personal_information_only
```

Where:

- `9b2b4508-aa2f-4940-b347-d1451example` is the unique identifier for the rule.
- `personal_information_only` is the new name for the rule.

If the command runs successfully, you receive output similar to the following.

```
{  
  "arn": "arn:aws:macie2:us-west-2:123456789012:findings-filter/9b2b4508-aa2f-4940-  
b347-d1451example",  
  "id": "9b2b4508-aa2f-4940-b347-d1451example"  
}
```

Where `arn` is the Amazon Resource Name (ARN) of the rule that was changed, and `id` is the unique identifier for the rule.

Similarly, the following example converts a suppression rule to a filter rule by changing the value for the action parameter from ARCHIVE to NOOP.

```
C:\> aws macie2 update-findings-filter --id 8a1c3508-aa2f-4940-b347-d1451example --  
action NOOP
```

Where:

- `8a1c3508-aa2f-4940-b347-d1451example` is the unique identifier for the rule.
- `NOOP` is the new action for Macie to perform on findings that meet the criteria of the rule—perform no action (don't suppress the findings).

If the command runs successfully, you receive output similar to the following:

```
{  
  "arn": "arn:aws:macie2:us-west-2:123456789012:findings-filter/8a1c3508-aa2f-4940-  
b347-d1451example",  
  "id": "8a1c3508-aa2f-4940-b347-d1451example"  
}
```

Where `arn` is the Amazon Resource Name (ARN) of the rule that was changed, and `id` is the unique identifier for the rule.


## Deleting filter rules

You can delete a filter rule at any time using the Amazon Macie console or the Amazon Macie API.

### Console

Follow these steps to delete a filter rule by using the Amazon Macie console.

#### To delete a filter rule

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Findings**.
3. In the **Saved rules** list, choose the edit icon () next to the filter rule that you want to delete.
4. Under **Filter rule**, choose **Delete**.

### AWS CLI

To delete a filter rule by using the AWS CLI, run the `delete-findings-filter` command. For the `id` parameter, specify the unique identifier for the filter rule to delete. You can get this identifier by running the `list-findings-filters` command to retrieve a list of filter and suppression rules for your account.

The following example deletes the filter rule whose unique identifier is `9b2b4508-aa2f-4940-b347-d1451example`.

```
C:\> aws macie2 delete-findings-filter --id 9b2b4508-aa2f-4940-b347-d1451example
```

If the command runs successfully, Macie returns an empty HTTP 200 response. Otherwise, Macie returns an HTTP 4xx or 500 response that indicates why the operation failed.

## Fields for filtering findings

To help you analyze findings more efficiently, the Amazon Macie console and the Amazon Macie API provide access to several sets of fields for filtering findings:

- **Common fields** – These fields store data that applies to any type of finding. They correlate to common attributes of findings such as severity, finding type, and finding ID.
- **Affected resource fields** – These fields store data about the resources that a finding applies to, such as the name, public access settings, and encryption settings for an affected S3 bucket or object.
- **Policy fields** – These fields store data that's specific to policy findings, such as the action that produced a finding, and the entity that performed the action.
- **Sensitive data classification fields** – These fields store data that's specific to sensitive data findings, such as the types of sensitive data that Macie found, and the unique identifier for the sensitive data discovery job that produced a finding.

A filter can use a combination of fields from any of the preceding sets.

The topics in this section list and describe the individual fields that you can use to filter findings. For additional details about these fields, including any relationships between the fields, see [Findings Descriptions](#) in the *Amazon Macie API Reference*.

### Topics

- [Common fields \(p. 134\)](#)
- [Affected resource fields \(p. 136\)](#)
- [Policy fields \(p. 141\)](#)
- [Sensitive data classification fields \(p. 147\)](#)

## Common fields

The following table lists and describes fields that you can use to filter findings based on common finding attributes. These fields store data that applies to any type of finding.

In the table, the **Field** column indicates the name of the field on the Amazon Macie console. The **JSON field** column uses dot notation to indicate the name of the field in JSON representations of findings and the Amazon Macie API. The **Description** column provides a brief description of the data that the field stores, and indicates any requirements for filter values. The table is sorted in ascending alphabetical order by field, and then by JSON field.

Field	JSON field	Description
Account ID*	accountId	The unique identifier for the Amazon Web Services account that the finding applies to. This is typically the account that owns the affected resource.
—	archived	A Boolean value that specifies whether the finding was archived by a suppression rule.  To add this field to a filter on the console, choose <b>Current</b> , <b>Archived</b> , or <b>All</b> in the filter bar.

Field	JSON field	Description
Category	category	<p>The category of the finding.</p> <p>The console provides a list of values to choose from when you add this field to a filter. In the API, valid values are: <code>CLASSIFICATION</code>, for a sensitive data finding; and, <code>POLICY</code>, for a policy finding.</p>
—	count	<p>The total number of occurrences of the finding. For sensitive data findings, this value is always 1. All sensitive data findings are considered unique because they derive from individual jobs.</p> <p>This field isn't available as a filter option on the console. With the API, you can use this field to define a numeric range for a filter.</p>
Created at	createdAt	<p>The date and time when Macie created the finding.</p> <p>You can use this field to define a time range for a filter.</p>
Finding ID*	id	<p>The unique identifier for the finding. This is a random string that Macie generates and assigns to a finding when it creates the finding.</p>
Finding type*	type	<p>The type of the finding—for example, <code>SensitiveData:S3Object/Personal</code> or <code>Policy:IAMUser/S3BucketPublic</code>.</p> <p>The console provides a list of values to choose from when you add this field to a filter. For a list of valid values in the API, see <a href="#">FindingType</a> in the <i>Amazon Macie API Reference</i>.</p>
Region	region	<p>The AWS Region that Macie created the finding in—for example, <code>us-east-1</code> or <code>ca-central-1</code>.</p>

Field	JSON field	Description
Sample	sample	<p>A Boolean value that specifies whether the finding is a sample finding. A <i>sample finding</i> is a finding that uses example data to demonstrate what a finding might contain.</p> <p>The console provides a list of values to choose from when you add this field to a filter.</p>
Severity	severity.description	<p>The qualitative representation of the finding's severity.</p> <p>The console provides a list of values to choose from when you add this field to a filter. In the API, valid values are: Low, Medium, and High.</p>
Updated at	updatedAt	<p>The date and time when the finding was last updated. For sensitive data findings, this value is the same as the value for the <b>Created at</b> field. All sensitive data findings are considered new because they derive from individual jobs.</p> <p>You can use this field to define a time range for a filter.</p>

\* To specify multiple values for this field on the console, add a condition that uses the field and specifies a distinct value for the filter, and then repeat that step for each additional value. To do this with the API, use an array that lists the values to use for the filter.

## Affected resource fields

The following topics list and describe the fields that you can use to filter findings based on the resource that a finding applies to. The topics are organized by resource type.

### Topics

- [S3 bucket \(p. 136\)](#)
- [S3 object \(p. 140\)](#)

## S3 bucket

The following table lists and describes fields that you can use to filter findings based on characteristics of the S3 bucket that a finding applies to.

In the table, the **Field** column indicates the name of the field on the Amazon Macie console. The **JSON field** column uses dot notation to indicate the name of the field in JSON representations of findings and the Amazon Macie API. (Longer JSON field names use the newline character sequence (\n) to improve readability.) The **Description** column provides a brief description of the data that the field stores, and

indicates any requirements for filter values. The table is sorted in ascending alphabetical order by field, and then by JSON field.

Field	JSON field	Description
—	<code>resourcesAffected.s3Bucket.createdTime</code>	<p>The date and time when the affected bucket was created.</p> <p>This field isn't available as a filter option on the console. With the API, you can use this field to define a time range for a filter.</p>
S3 bucket default encryption	<code>resourcesAffected.s3Bucket.typeOfServerSideEncryption</code>	<p>The type of server-side encryption that's used by default to encrypt objects that are added to the affected bucket.</p> <p>The console provides a list of values to choose from when you add this field to a filter. For a list of valid values for the API, see <a href="#">EncryptionType</a> in the <i>Amazon Macie API Reference</i>.</p>
S3 bucket encryption KMS master key id*	<code>resourcesAffected.s3Bucket.kmsMasterKeyArn</code>	<p>The Amazon Resource Name (ARN) or unique identifier (key ID) for the AWS KMS key that's used by default to encrypt objects that are added to the affected bucket.</p>
S3 bucket encryption required by bucket policy	<code>resourcesAffected.s3Bucket.bucketPolicyRequiresEncryption</code>	<p>Specifies whether the bucket policy for the affected bucket requires server-side encryption of objects when objects are uploaded to the bucket.</p> <p>The console provides a list of values to choose from when you add this field to a filter. For a list of valid values for the API, see <a href="#">S3Bucket</a> in the <i>Amazon Macie API Reference</i>.</p>
S3 bucket name*	<code>resourcesAffected.s3Bucket.name</code>	<p>The name of the affected bucket.</p>
S3 bucket owner display name*	<code>resourcesAffected.s3Bucket.ownerDisplayName</code>	<p>The display name of the AWS user who owns the affected bucket.</p>
S3 bucket public access permission	<code>resourcesAffected.s3Bucket.publicAccessPermission</code>	<p>Specifies whether the affected bucket is publicly accessible based on a combination of permissions settings that apply to the bucket.</p>

Field	JSON field	Description
—	resourcesAffected.s3BucketPublicAccessConfiguration.accountLevelPermissions.blockPublicAccess.blockPublicAcls	<p>The console provides a list of values to choose from when you add this field to a filter. For a list of valid values for the API, see <a href="#">BucketPublicAccess</a> in the <i>Amazon Macie API Reference</i>.</p> <p>A Boolean value that specifies whether Amazon S3 blocks public access control lists (ACLs) for the affected bucket and objects in the bucket. This is an account-level, block public access setting for the bucket.</p> <p>This field isn't available as a filter option on the console.</p>
—	resourcesAffected.s3BucketPublicAccessConfiguration.accountLevelPermissions.blockPublicAccess.blockPublicPolicy	<p>A Boolean value that specifies whether Amazon S3 blocks public bucket policies for the affected bucket. This is an account-level, block public access setting for the bucket.</p> <p>This field isn't available as a filter option on the console.</p>
—	resourcesAffected.s3BucketPublicAccessConfiguration.accountLevelPermissions.blockPublicAccess.ignorePublicAcls	<p>A Boolean value that specifies whether Amazon S3 ignores public ACLs for the affected bucket and objects in the bucket. This is an account-level, block public access setting for the bucket.</p> <p>This field isn't available as a filter option on the console.</p>
—	resourcesAffected.s3BucketPublicAccessConfiguration.accountLevelPermissions.restrictPublicBucket	<p>A Boolean value that specifies whether Amazon S3 restricts public bucket policies for the affected bucket. This is an account-level, block public access setting for the bucket.</p> <p>This field isn't available as a filter option on the console.</p>
—	resourcesAffected.s3BucketPublicAccessConfiguration.bucketLevelPermissions.allowPublicReadAccess	<p>A Boolean value that specifies whether the bucket-level ACL for the affected bucket grants the general public with read access permissions for the bucket.</p> <p>This field isn't available as a filter option on the console.</p>

Field	JSON field	Description
—	<code>resourcesAffected.s3Bucket</code> <code>bucketLevelPermissions.accessControlListWithWritePublicWriteAccessPermissions</code>	A Boolean value that specifies whether the bucket-level ACL for the affected bucket grants the specified public write access permissions for the bucket.  This field isn't available as a filter option on the console.
—	<code>resourcesAffected.s3Bucket</code> <code>bucketLevelPermissions.blockPublicAccess</code>	A Boolean value that specifies whether Amazon S3 blocks public ACLs for the affected bucket and objects in the bucket.  This is a bucket-level, block public access setting for a bucket.  This field isn't available as a filter option on the console.
—	<code>resourcesAffected.s3Bucket</code> <code>bucketLevelPermissions.blockPublicPolicy</code>	A Boolean value that specifies whether Amazon S3 blocks public bucket policies for the affected bucket. This is a bucket-level, block public access setting for the bucket.  This field isn't available as a filter option on the console.
—	<code>resourcesAffected.s3Bucket</code> <code>bucketLevelPermissions.blockPublicAcls</code>	A Boolean value that specifies whether Amazon S3 ignores public ACLs for the affected bucket and objects in the bucket.  This is a bucket-level, block public access setting for the bucket.  This field isn't available as a filter option on the console.
—	<code>resourcesAffected.s3Bucket</code> <code>bucketLevelPermissions.blockPublicBuckets</code>	A Boolean value that specifies whether Amazon S3 restricts public bucket policies for the affected bucket. This is a bucket-level, block public access setting for the bucket.  This field isn't available as a filter option on the console.



Field	JSON field	Description
—	<code>resourcesAffected.s3Bucket</code> <code>bucketLevelPermissions.bucketPolicyAllowsPublicReadAccess</code>	A Boolean value that specifies whether the affected bucket's policy allows the general public to have read access to the bucket.  This field isn't available as a filter option on the console.
—	<code>resourcesAffected.s3Bucket</code> <code>bucketLevelPermissions.bucketPolicyAllowsPublicWriteAccess</code>	A Boolean value that specifies whether the affected bucket's policy allows the general public to have write access to the bucket.  This field isn't available as a filter option on the console.
S3 bucket tag key*	<code>resourcesAffected.s3Bucket</code>	A tag key that's associated with the affected bucket.
S3 bucket tag value*	<code>resourcesAffected.s3Bucket</code>	A tag value that's associated with the affected bucket.

\* To specify multiple values for this field on the console, add a condition that uses the field and specifies a distinct value for the filter, and then repeat that step for each additional value. To do this with the API, use an array that lists the values to use for the filter.

## S3 object

The following table lists and describes fields that you can use to filter findings based on characteristics of the S3 object that a finding applies to.

In the table, the **Field** column indicates the name of the field on the Amazon Macie console. The **JSON field** column uses dot notation to indicate the name of the field in JSON representations of findings and the Amazon Macie API. The **Description** column provides a brief description of the data that the field stores, and indicates any requirements for filter values. The table is sorted in ascending alphabetical order by field, and then by JSON field.

Field	JSON field	Description
S3 object encryption KMS master key id*	<code>resourcesAffected.s3Object</code>	The Amazon Resource Name (ARN) or unique identifier (key ID) for the AWS KMS key that was used to encrypt the affected object.
S3 object encryption type	<code>resourcesAffected.s3Object</code>	The type of server-side encryption that was used to encrypt the affected object.  The console provides a list of values to choose from when you add this field to a filter. For a list of valid values for the API, see

Field	JSON field	Description
		<a href="#">EncryptionType</a> in the <i>Amazon Macie API Reference</i> .
—	<code>resourcesAffected.s3Object.fileName</code>	The file name extension of the affected object. For objects that don't have a file name extension, specify "" as the value for the filter.  This field isn't available as a filter option on the console.
—	<code>resourcesAffected.s3Object.lastModifiedDate</code>	The date and time when the affected object was created or last changed, whichever is latest.  This field isn't available as a filter option on the console. With the API, you can use this field to define a time range for a filter.
S3 object key*	<code>resourcesAffected.s3Object.key</code>	The full key (name) that's assigned to the affected object.
—	<code>resourcesAffected.s3Object.path</code>	The path to the affected object, including the full key (name).  This field isn't available as a filter option on the console.
S3 object public access	<code>resourcesAffected.s3Object.publicAccess</code>	A Boolean value that specifies whether the affected object is publicly accessible based on a combination of permission settings that apply to the object.  The console provides a list of values to choose from when you add this field to a filter.
S3 object tag key*	<code>resourcesAffected.s3Object.tagKey</code>	A tag key that's associated with the affected object.
S3 object tag value*	<code>resourcesAffected.s3Object.tagValue</code>	A tag value that's associated with the affected object.

\* To specify multiple values for this field on the console, add a condition that uses the field and specifies a distinct value for the filter, and then repeat that step for each additional value. To do this with the API, use an array that lists the values to use for the filter.

## Policy fields

The following table lists and describes fields that you can use to filter policy findings. These fields store data that's specific to policy findings.

In the table, the **Field** column indicates the name of the field on the Amazon Macie console. The **JSON field** column uses dot notation to indicate the name of the field in JSON representations of findings and the Amazon Macie API. (Longer JSON field names use the newline character sequence (\n) to improve readability.) The **Description** column provides a brief description of the data that the field stores, and indicates any requirements for filter values. The table is sorted in ascending alphabetical order by field, and then by JSON field.

Field	JSON field	Description
Action type	policyDetails.action.actionType	The type of action that produced the finding. The only valid value for this field is <code>AWS_API_CALL</code> .
API call name*	policyDetails.action.apiCallName	The name of the operation that was invoked most recently and produced the finding—for example, <code>DeleteBucketEncryption</code> .
API service name*	policyDetails.action.apiCallServiceName	The URL of the AWS service that provides the operation that was invoked and produced the finding—for example, <code>s3.amazonaws.com</code> .
—	policyDetails.action.apiCallTimestamp	The first date and time when any operation was invoked and produced the finding.  This field isn't available as a filter option on the console. With the API, you can use this field to define a time range for a filter.
—	policyDetails.action.apiCallTimestampMs	The most recent date and time when the specified operation ( <b>API call name</b> or <code>api</code> ) was invoked and produced the finding.  This field isn't available as a filter option on the console. With the API, you can use this field to define a time range for a filter.
—	policyDetails.actor.domainName	The domain name of the device that was used to perform the action.  This field isn't available as a filter option on the console.
IP city*	policyDetails.actor.ipAddressCity	The name of the originating city for the IP address of the device that was used to perform the action.

Field	JSON field	Description
IP country*	<code>policyDetails.actor.ipAddress</code>	The name of the originating country for the IP address of the device that was used to perform the action—for example, United States.
—	<code>policyDetails.actor.ipAddress</code>	The Autonomous System asn Number (ASN) for the autonomous system that included the IP address of the device that was used to perform the action.  This field isn't available as a filter option on the console.
IP owner ASN org*	<code>policyDetails.actor.ipAddress</code>	The organization identifies that associated with the ASN for the autonomous system that included the IP address of the device that was used to perform the action.
IP owner ISP*	<code>policyDetails.actor.ipAddress</code>	The name of the Internet service provider (ISP) that owned the IP address of the device that was used to perform the action.
IP V4 address*	<code>policyDetails.actor.ipAddress</code>	The Internet Protocol version 4 (IPv4) address of the device that was used to perform the action.
—	<code>policyDetails.actor.userId</code>	For an action performed with temporary security credentials that were obtained using the AssumeRole operation of the AWS STS API, the AWS access key ID that identifies the credentials.  This field isn't available as a filter option on the console.
User identity assumed role account id*	<code>policyDetails.actor.userId</code>	For an action performed with temporary security credentials that were obtained using the AssumeRole operation of the AWS STS API, the unique identifier for the Amazon Web Services account that owns the entity that was used to get the credentials.

Field	JSON field	Description
User identity assumed role principal id*	<code>policyDetails.actor.userIdentityAssumedRolePrincipalId</code>	For an action performed with temporary security credentials that were obtained using the AssumeRole operation of the AWS STS API, the unique identifier for the entity that was used to get the credentials.
User identity assumed role session ARN*	<code>policyDetails.actor.userIdentityAssumedRoleSessionArn</code>	For an action performed with temporary security credentials that were obtained using the AssumeRole operation of the AWS STS API, the Amazon Resource Name (ARN) of the source account, IAM user, or role that was used to get the credentials.
—	<code>policyDetails.actor.userIdentityAssumedRoleSessionContext.type</code>	For an action performed with temporary security credentials that were obtained using the AssumeRole operation of the AWS STS API, the source of the temporary security credentials—for example, Root, IAMUser, or Role.  This field isn't available as a filter option on the console.
—	<code>policyDetails.actor.userIdentityAssumedRoleSessionContext.userName</code>	For an action performed with temporary security credentials that were obtained using the AssumeRole operation of the AWS STS API, the name or alias of the user or role that issued the session. Note that this value is null if the credentials were obtained from a root account that doesn't have an alias.  This field isn't available as a filter option on the console.
User identity AWS account account id*	<code>policyDetails.actor.userIdentityAssumedRoleSourceAccountId</code>	For an action performed using the credentials for another Amazon Web Services account, the unique identifier for the account.
User identity AWS account principal id*	<code>policyDetails.actor.userIdentityAssumedRoleSourcePrincipalId</code>	For an action performed using the credentials for another Amazon Web Services account, the unique identifier for the entity that performed the action.

Field	JSON field	Description
User identity AWS service invoked by	<code>policyDetails.actor.userIdentityDetails.serviceName</code>	For any actions performed by an account that belongs to an Amazon Web Services service, the name of the service.
—	<code>policyDetails.actor.userIdentityDetails.accessKeyId</code>	For any actions performed with temporary security credentials that were obtained using the <code>GetFederationToken</code> operation of the AWS STS API, the AWS access key ID that identifies the credentials.  This field isn't available as a filter option on the console.
User identity federated session ARN*	<code>policyDetails.actor.userIdentityDetails.sessionArn</code>	For any actions performed with temporary security credentials that were obtained using the <code>GetFederationToken</code> operation of the AWS STS API, the ARN of the entity that was used to get the credentials.
User identity federated user account id*	<code>policyDetails.actor.userIdentityDetails.accountId</code>	For any actions performed with temporary security credentials that were obtained using the <code>GetFederationToken</code> operation of the AWS STS API, the unique identifier for the Amazon Web Services account that owns the entity that was used to get the credentials.
User identity federated user principal id*	<code>policyDetails.actor.userIdentityDetails.principalId</code>	For any actions performed with temporary security credentials that were obtained using the <code>GetFederationToken</code> operation of the AWS STS API, the unique identifier for the entity that was used to get the credentials.
—	<code>policyDetails.actor.userIdentityDetails.sessionContext.sessionIssuer.type</code>	For any actions performed with temporary security credentials that were obtained using the <code>GetFederationToken</code> operation of the AWS STS API, the source of the temporary security credentials—for example, <code>Root</code> , <code>IAMUser</code> , or <code>Role</code> .  This field isn't available as a filter option on the console.

Field	JSON field	Description
—	<code>policyDetails.actor.userName</code> <code>sessionIssuer.userName</code>	For an action performed with temporary security credentials that were obtained using the <code>GetFederationToken</code> operation of the AWS STS API, the name or alias of the user or role that issued the session. Note that this value is null if the credentials were obtained from a root account that doesn't have an alias.  This field isn't available as a filter option on the console.
User identity IAM account id*	<code>policyDetails.actor.userId</code>	For an action performed using an IAM user's credentials, the unique identifier for the Amazon Web Services account that's associated with the IAM user who performed the action.
User identity IAM principal id*	<code>policyDetails.actor.userId</code>	For an action performed using an IAM user's credentials, the unique identifier for the IAM user who performed the action.
User identity IAM user name*	<code>policyDetails.actor.userId</code>	For an action performed using an IAM user's credentials, the user name of the IAM user who performed the action.
User identity root account id*	<code>policyDetails.actor.userId</code>	For an action performed using the credentials for your Amazon Web Services account, the unique identifier for the account.
User identity root principal id*	<code>policyDetails.actor.userId</code>	For an action performed using the credentials for your Amazon Web Services account, the unique identifier for the entity that performed the action.
User identity type	<code>policyDetails.actor.userId</code>	The type of entity that performed the action that produced the finding.  The console provides a list of values to choose from when you add this field to a filter. For a list of valid values for the API, see <a href="#">UserIdentityType</a> in the <i>Amazon Macie API Reference</i> .

\* To specify multiple values for this field on the console, add a condition that uses the field and specifies a distinct value for the filter, and then repeat that step for each additional value. To do this with the API, use an array that lists the values to use for the filter.

## Sensitive data classification fields

The following table lists and describes fields that you can use to filter sensitive data findings. These fields store data that's specific to sensitive data findings.

In the table, the **Field** column indicates the name of the field on the Amazon Macie console. The **JSON field** column uses dot notation to indicate the name of the field in JSON representations of findings and the Amazon Macie API. The **Description** column provides a brief description of the data that the field stores, and indicates any requirements for filter values. The table is sorted in ascending alphabetical order by field, and then by JSON field.

Field	JSON field	Description
Customer data identifier detection ARN*	classificationDetails.results.detections.customerDataIdentifierArn	The Amazon Resource Name (ARN) of the custom data identifier that detected the data and produced the finding.
Customer data identifier detection name*	classificationDetails.results.detections.customerDataIdentifierName	The name of the custom data identifier that detected the data and produced the finding.
Customer data identifier total count	classificationDetails.results.detections.customerDataIdentifierTotalCount	The total number of occurrences of data that was detected by custom data identifiers and produced the finding.  You can use this field to define a numeric range for a filter.
Job ID*	classificationDetails.jobId	The unique identifier for the sensitive data discovery job that produced the finding.
—	classificationDetails.results.detections.contentType	The type of content, as a MIME type, that the finding applies to—for example, text/csv for a CSV file or application/pdf for an Adobe Portable Document Format file.  This field isn't available as a filter option on the console.
—	classificationDetails.results.detections.s3ObjectSize	The total storage size, in bytes, of the S3 object that the finding applies to.  This field isn't available as a filter option on the console. With the API, you can use this field to define a numeric range for a filter.



Field	JSON field	Description
Result status code*	classificationDetails.resultStatus	<p>The status of the finding. Valid values are:</p> <ul style="list-style-type: none"> <li>COMPLETE – Macie completed its analysis of the object.</li> <li>PARTIAL – Macie analyzed only a subset of the data in the object. For example, the object is an archive file that contains files in an unsupported format.</li> <li>SKIPPED – Macie wasn't able to analyze the object. For example, the object is a malformed file.</li> </ul>
Sensitive data category	classificationDetails.resultSensitiveDataCategory	<p>The category of sensitive data that was detected and produced the finding.</p> <p>The console provides a list of values to choose from when you add this field to a filter. In the API, valid values are: CREDENTIALS, FINANCIAL_INFORMATION, and PERSONAL_INFORMATION.</p>
Sensitive data detection type	classificationDetails.resultSensitiveDataDetectionType	<p>The type of sensitive data that was detected and produced the finding.</p> <p>The console provides a list of values to choose from when you add this field to a filter. For a complete list of types, see <a href="#">Sensitive data detection types (p. 148)</a>.</p>
Sensitive data total count	classificationDetails.resultSensitiveDataOccurrencesCount	<p>The total number of occurrences of the sensitive data that was detected and produced the finding.</p> <p>You can use this field to define a numeric range for a filter.</p>

\* To specify multiple values for this field on the console, add a condition that uses the field and specifies a distinct value for the filter, and then repeat that step for each additional value. To do this with the API, use an array that lists the values to use for the filter.

## Sensitive data detection types

The following topics list values that you can specify for the **Sensitive data detection type** field in a filter. (The JSON name of this field is

`classificationDetails.result.sensitiveData.detections.type`.) The topics are organized by the categories of sensitive data that Macie can detect using managed data identifiers.

### Categories

- [Credentials](#) (p. 149)
- [Financial information](#) (p. 149)
- [Personal information](#) (p. 150)

To learn more about a specific detection type, see [Using managed data identifiers](#) (p. 37).

### Credentials

You can specify the following values to filter findings that report occurrences of credentials data in S3 objects.

Detection type	Filter values
AWS secret key	AWS_CREDENTIALS
OpenSSH private key	OPENSSSH_PRIVATE_KEY
PGP private key	PGP_PRIVATE_KEY
Public Key Cryptography Standard (PKCS) private key	PKCS
PuTTY private key	PUTTY_PRIVATE_KEY

### Financial information

You can specify the following values to filter findings that report occurrences of financial information in S3 objects.

Detection type	Filter values
Bank account number	BANK_ACCOUNT_NUMBER (for Canadian and US bank account numbers), FRANCE_BANK_ACCOUNT_NUMBER, GERMANY_BANK_ACCOUNT_NUMBER, ITALY_BANK_ACCOUNT_NUMBER, SPAIN_BANK_ACCOUNT_NUMBER, UK_BANK_ACCOUNT_NUMBER
Credit card expiration date	CREDIT_CARD_EXPIRATION
Credit card magnetic strip data	CREDIT_CARD_MAGNETIC_STRIPE
Credit card number	CREDIT_CARD_NUMBER (for credit card numbers that are in proximity of a keyword) and CREDIT_CARD_NUMBER_(NO_KEYWORD) (for credit card numbers that aren't in proximity of a keyword)
Credit card verification code	CREDIT_CARD_SECURITY_CODE

## Personal information

You can specify the following values to filter findings that report occurrences of personal health information (PHI) in S3 objects.

Detection type	Filter values
Drug Enforcement Agency (DEA) Registration Number	US_DRUG_ENFORCEMENT_AGENCY_NUMBER
Health Insurance Claim Number (HICN)	USA_HEALTH_INSURANCE_CLAIM_NUMBER
Health insurance or medical identification number	CANADA_HEALTH_NUMBER, EUROPEAN_HEALTH_INSURANCE_CARD_NUMBER, FINLAND_EUROPEAN_HEALTH_INSURANCE_NUMBER, FRANCE_HEALTH_INSURANCE_NUMBER, UK_NHS_NUMBER, USA_MEDICARE_BENEFICIARY_IDENTIFIER
Healthcare Common Procedure Coding System (HCPCS) code	USA_HEALTHCARE_PROCEDURE_CODE
National Drug Code (NDC)	USA_NATIONAL_DRUG_CODE
National Provider Identifier (NPI)	USA_NATIONAL_PROVIDER_IDENTIFIER
Unique device identifier (UDI)	MEDICAL_DEVICE_UDI

You can specify the following values to filter findings that report occurrences of personally identifiable information (PII) in S3 objects.

Detection type	Filter values
Birth date	DATE_OF_BIRTH
Driver's license identification number	AUSTRALIA_DRIVERS_LICENSE, AUSTRIA_DRIVERS_LICENSE, BELGIUM_DRIVERS_LICENSE, BULGARIA_DRIVERS_LICENSE, CANADA_DRIVERS_LICENSE, CROATIA_DRIVERS_LICENSE, CYPRUS_DRIVERS_LICENSE, CZECHIA_DRIVERS_LICENSE, DENMARK_DRIVERS_LICENSE, DRIVERS_LICENSE (for the US), ESTONIA_DRIVERS_LICENSE, FINLAND_DRIVERS_LICENSE, FRANCE_DRIVERS_LICENSE, GERMANY_DRIVERS_LICENSE, GREECE_DRIVERS_LICENSE, HUNGARY_DRIVERS_LICENSE, IRELAND_DRIVERS_LICENSE, ITALY_DRIVERS_LICENSE, LATVIA_DRIVERS_LICENSE, LITHUANIA_DRIVERS_LICENSE, LUXEMBOURG_DRIVERS_LICENSE, MALTA_DRIVERS_LICENSE,

Detection type	Filter values
	NETHERLANDS_DRIVERS_LICENSE, POLAND_DRIVERS_LICENSE, PORTUGAL_DRIVERS_LICENSE, ROMANIA_DRIVERS_LICENSE, SLOVAKIA_DRIVERS_LICENSE, SLOVENIA_DRIVERS_LICENSE, SPAIN_DRIVERS_LICENSE, SWEDEN_DRIVERS_LICENSE, UK_DRIVERS_LICENSE
Electoral roll number	UK_ELECTORAL_ROLL_NUMBER
Full name	NAME
Global Positioning System (GPS) coordinates	LATITUDE_LONGITUDE
Mailing address	ADDRESS, BRAZIL_CEP_CODE
National identification number	BRAZIL_RG_NUMBER, FRANCE_NATIONAL_IDENTIFICATION_NUMBER, GERMANY_NATIONAL_IDENTIFICATION_NUMBER, ITALY_NATIONAL_IDENTIFICATION_NUMBER, SPAIN_DNI_NUMBER
National Insurance Number (NINO)	UK_NATIONAL_INSURANCE_NUMBER
Passport number	CANADA_PASSPORT_NUMBER, FRANCE_PASSPORT_NUMBER, GERMANY_PASSPORT_NUMBER, ITALY_PASSPORT_NUMBER, SPAIN_PASSPORT_NUMBER, UK_PASSPORT_NUMBER, USA_PASSPORT_NUMBER
Permanent residence number	CANADA_NATIONAL_IDENTIFICATION_NUMBER
Phone number	BRAZIL_PHONE_NUMBER, FRANCE_PHONE_NUMBER, GERMANY_PHONE_NUMBER, ITALY_PHONE_NUMBER, PHONE_NUMBER (for Canada and the US), SPAIN_PHONE_NUMBER, UK_PHONE_NUMBER
Social Insurance Number (SIN)	CANADA_SOCIAL_INSURANCE_NUMBER
Social Security number (SSN)	SPAIN_SOCIAL_SECURITY_NUMBER, USA_SOCIAL_SECURITY_NUMBER
Taxpayer identification or reference number	AUSTRALIA_TAX_FILE_NUMBER, BRAZIL_CNPJ_NUMBER, BRAZIL_CPF_NUMBER, FRANCE_TAX_IDENTIFICATION_NUMBER, GERMANY_TAX_IDENTIFICATION_NUMBER, SPAIN_NIE_NUMBER, SPAIN_NIF_NUMBER, SPAIN_TAX_IDENTIFICATION_NUMBER, UK_TAX_IDENTIFICATION_NUMBER, USA_INDIVIDUAL_TAX_IDENTIFICATION_NUMBER
Vehicle identification number (VIN)	VEHICLE_IDENTIFICATION_NUMBER

# Suppressing Amazon Macie findings

To streamline your analysis of findings, you can create and use *suppression rules*. A *suppression rule* is a set of attribute-based filter criteria that defines cases where you want Amazon Macie to archive findings automatically. Suppression rules are helpful in situations where you've reviewed a class of findings and don't want to be notified of them again.

For example, you might decide to allow S3 buckets to contain mailing addresses, if the buckets don't allow public access and they encrypt new objects by default. In this case, you can create a suppression rule that specifies filter criteria for the following fields: **Sensitive data detection type**, **S3 bucket public access permission**, and **S3 bucket default encryption**. The rule suppresses future findings that meet the filter criteria.

If you suppress findings by using a suppression rule, Macie continues to generate findings for subsequent occurrences of sensitive data and potential policy violations that meet the rule's criteria. However, Macie automatically changes the status of the findings to *archived*. This means that the findings don't appear by default on the Amazon Macie console, but they persist in Macie until they expire. (Macie stores your findings for 90 days.)

In addition to changing the status of suppressed findings, Macie doesn't publish the findings to Amazon EventBridge as events (formerly called Amazon CloudWatch Events) or to AWS Security Hub. Macie does, however, continue to create and store [sensitive data discovery results \(p. 80\)](#) that correlate to sensitive data findings that you suppress. This helps ensure that you have an immutable history of sensitive data findings for data privacy and protection audits or investigations that you perform.

To create and manage suppression rules, you can use the Amazon Macie console or the Amazon Macie API. The following topics explain how. For the API, the topics explain how to perform these tasks with the [AWS Command Line Interface \(AWS CLI\)](#). You can also perform these tasks by sending HTTPS requests directly to Macie, or by using a current version of another AWS command line tool or an AWS SDK. For information about AWS tools and SDKs, see [Tools to Build on AWS](#).

## Topics

- [Creating suppression rules \(p. 152\)](#)
- [Viewing suppressed findings \(p. 154\)](#)
- [Changing suppression rules \(p. 155\)](#)
- [Deleting suppression rules \(p. 156\)](#)

## Creating suppression rules

Before you create a suppression rule, it's important to note that you can't restore (unarchive) findings that you suppress using a suppression rule. You can, however, [view suppressed findings \(p. 154\)](#) on the Amazon Macie console and access suppressed findings with the Amazon Macie API.

When you create a suppression rule, you specify filter criteria, a name, and, optionally, a description for the rule. You can create a suppression rule using the Amazon Macie console or the Amazon Macie API.

### Console

Follow these steps to create a suppression rule by using the Amazon Macie console.

#### To create a suppression rule

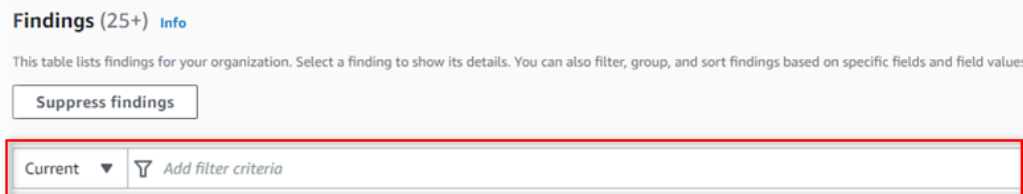
1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Findings**.

### Tip

To use an existing suppression or filter rule as a starting point, choose the rule from the **Saved rules** list.

You can also streamline creation of a rule by first pivoting and drilling down on findings by a predefined logical group. If you do this, Macie automatically creates and applies the appropriate filter conditions, which can be a helpful starting point for creating a rule. To do this, choose **By bucket**, **By type**, or **By job** in the navigation pane (under **Findings**), and then choose an item in the table. In the details panel, choose the link for the field to pivot on.

3. In the filter bar, add filter conditions that specify attributes of the findings that you want the rule to suppress.



To learn how to add filter conditions, see [Creating and applying filters to findings \(p. 123\)](#).

4. When you finish adding filter conditions for the rule, choose **Suppress findings** above the filter bar.
5. Under **Suppression rule**, enter a name and, optionally, a description for the rule.
6. Choose **Save**.

## AWS CLI

To create a suppression rule by using the AWS CLI, run the `create-findings-filter` command and specify the appropriate values for the required parameters. For the `action` parameter, specify `ARCHIVE` to ensure that Macie suppresses findings that meet the criteria of the rule.

For the `criterion` parameter, specify a map of conditions that define the filter criteria for the rule. In the map, each condition should specify a field, an operator, and one or more values for the field. The type and number of values depends on the field and operator that you choose. For information about the fields, operators, and types of values that you can use in a condition, see [Fields for filtering findings \(p. 134\)](#), [Using operators in conditions \(p. 120\)](#), and [Specifying values for fields \(p. 118\)](#).

The following examples create a suppression rule that returns all sensitive data findings that are in the current AWS Region and report occurrences of mailing addresses (and no other types of sensitive data) in S3 objects.

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

```
$ aws macie2 create-findings-filter \
--action ARCHIVE \
--name my_suppression_rule \
--finding-criteria '{"criterion":
{"classificationDetails.result.sensitiveData.detections.type":{"eqExactMatch":
["ADDRESS"]}}}'
```

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

```
C:\> aws macie2 create-findings-filter ^
--action ARCHIVE ^
--name my_suppression_rule ^
--finding-criteria={"criterion":
{"classificationDetails.result.sensitiveData.detections.type":{"eqExactMatch":
{"ADDRESS"}}}}
```

Where:

- `my_suppression_rule` is the custom name for the rule.
- `criterion` is a map of filter conditions for the rule:
  - `classificationDetails.result.sensitiveData.detections.type` is the JSON name of the **Sensitive data detection type** field.
  - `eqExactMatch` specifies the *equals exact match* operator.
  - `ADDRESS` is an enumerated value for the **Sensitive data detection type** field.

If the command runs successfully, you receive output similar to the following.

```
{
  "arn": "arn:aws:macie2:us-west-2:123456789012:findings-filter/8a3c5608-aa2f-4940-
b347-d1451example",
  "id": "8a3c5608-aa2f-4940-b347-d1451example"
}
```

Where `arn` is the Amazon Resource Name (ARN) of the suppression rule that was created, and `id` is the unique identifier for the rule.

For additional examples of filter criteria, see [Filtering findings programmatically with the Amazon Macie API](#) (p. 126).

## Viewing suppressed findings

By default, Macie doesn't display suppressed findings on the Amazon Macie console. However, you can view these findings on the console by changing your filter settings. The following procedure explains how.

### To view suppressed findings on the console

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Findings**. The **Findings** page displays current findings for your account in the current AWS Region. By default, this doesn't include findings that were suppressed by a suppression rule.
3. In the filter bar, do one of the following:
  - To display only suppressed findings, choose **Current**, and then choose **Archived**.
  - To display both suppressed and current findings, choose **Current**, and then choose **All**.

You can also access suppressed findings by using the Amazon Macie API. To retrieve a list of suppressed findings, use the [ListFindings](#) operation and include a filter condition that specifies `true` for the `archived` field. For an example of how to do this using the AWS CLI, see [Filtering findings programmatically](#) (p. 126). To retrieve the details of one or more suppressed findings, use the [GetFindings](#) operation and specify the unique identifier for each finding to retrieve.

## Changing suppression rules

You can change the settings for a suppression rule at any time using the Amazon Macie console or the Amazon Macie API.

### Console

Follow these steps to change the settings for an existing suppression rule by using the Amazon Macie console.

#### To change a suppression rule

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Findings**.
3. In the **Saved rules** list, choose the edit icon (✎) next to the suppression rule that you want to change.
4. Do any of the following:
  - To change the name of the rule, enter a new name in the **Name** box under **Suppression rule**.
  - To change the description of the rule, enter a new description in the **Description** box under **Suppression rule**.
  - To change the filter criteria of the rule, use the filter bar to enter conditions that specify attributes of the findings that you want the rule to suppress. To learn how, see [Creating and applying filters to findings \(p. 123\)](#).
5. When you finish making changes, choose **Save**.

### AWS CLI

To change a suppression rule by using the AWS CLI, run the `update-findings-filter` command and use the supported parameters to specify a new value for each setting that you want to change. For the `id` parameter, specify the unique identifier for the rule to change. You can get this identifier by running the `list-findings-filters` command to retrieve a list of suppression and filter rules for your account.

The following example changes the name of an existing suppression rule.

```
C:\> aws macie2 update-findings-filter --id 8a3c5608-aa2f-4940-b347-d1451example --name mailing_addresses_only
```

Where:

- `8a3c5608-aa2f-4940-b347-d1451example` is the unique identifier for the rule.
- `mailing_addresses_only` is the new name for the rule.

If the command runs successfully, you receive output similar to the following.

```
{
  "arn": "arn:aws:macie2:us-west-2:123456789012:findings-filter/8a3c5608-aa2f-4940-b347-d1451example",
  "id": "8a3c5608-aa2f-4940-b347-d1451example"
}
```

Where `arn` is the Amazon Resource Name (ARN) of the rule that was changed, and `id` is the unique identifier for the rule.



Similarly, the following example converts a filter rule to a suppression rule by changing the value for the `action` parameter from `NOOP` to `ARCHIVE`.

```
C:\> aws macie2 update-findings-filter --id 8a1c3508-aa2f-4940-b347-d1451example --  
action ARCHIVE
```

Where:

- `8a1c3508-aa2f-4940-b347-d1451example` is the unique identifier for the rule.
- `ARCHIVE` is the new action for Macie to perform on findings that meet the criteria of the rule—suppress the findings.

If the command runs successfully, you receive output similar to the following:

```
{  
  "arn": "arn:aws:macie2:us-west-2:123456789012:findings-filter/8a1c3508-aa2f-4940-  
b347-d1451example",  
  "id": "8a1c3508-aa2f-4940-b347-d1451example"  
}
```

Where `arn` is the Amazon Resource Name (ARN) of the rule that was changed, and `id` is the unique identifier for the rule.

## Deleting suppression rules

You can delete a suppression rule at any time using the Amazon Macie console or the Amazon Macie API. If you delete a suppression rule, Macie stops suppressing new and subsequent occurrences of findings that meet the criteria of the rule and aren't suppressed by other rules. Note, however, that Macie might continue to suppress findings that it's currently processing and meet the rule's criteria.

After you delete a suppression rule, new and subsequent occurrences of findings that met the rule's criteria have a status of *current*. This means that they appear by default on the Amazon Macie console. In addition, Macie publishes these findings as Amazon EventBridge events. For policy findings, Macie also publishes the findings to AWS Security Hub.

### Console

Follow these steps to delete a suppression rule by using the Amazon Macie console.

#### To delete a suppression rule

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Findings**.
3. In the **Saved rules** list, choose the edit icon (✎) next to the suppression rule that you want to delete.
4. Under **Suppression rule**, choose **Delete**.

### AWS CLI

To delete a suppression rule by using the AWS CLI, run the `delete-findings-filter` command. For the `id` parameter, specify the unique identifier for the suppression rule to delete. You can get this identifier by running the `list-findings-filters` command to retrieve a list of suppression and filter rules for your account.

The following example deletes the suppression rule whose unique identifier is `8a3c5608-aa2f-4940-b347-d1451example`.

```
C:\> aws macie2 delete-findings-filter --id 8a3c5608-aa2f-4940-b347-d1451example
```

If the command runs successfully, Macie returns an empty HTTP 200 response. Otherwise, Macie returns an HTTP 4xx or 500 response that indicates why the operation failed.

## Severity scoring for Amazon Macie findings

When Amazon Macie generates a policy or sensitive data finding, it automatically assigns a severity to the finding. A finding's severity reflects the principal characteristics of the finding and can therefore help you assess and prioritize your findings. A finding's severity doesn't imply or otherwise indicate the criticality or importance that an affected resource might have for your organization.

For policy findings, severity is based on the nature of a potential issue with the security or privacy of your Amazon S3 data. For sensitive data findings, severity is based on the nature and number of occurrences of sensitive data that Macie found in an S3 object.

In Macie, a finding's severity is represented in two ways.

### Severity level

This is a qualitative representation of severity. Severity levels range from *Low*, for least severe, to *High*, for most severe.

Severity levels appear directly on the Amazon Macie console. They're also available in JSON representations of findings on the Macie console, the Amazon Macie API, and sensitive data discovery results that correlate to sensitive data findings. Severity levels are also included in finding events that Macie publishes to Amazon EventBridge and findings that Macie publishes to AWS Security Hub.

### Severity score

This is a numerical representation of severity. Severity scores range from 1 through 3 and map directly to severity levels:

Severity score	Severity level
1	Low
2	Medium
3	High

Severity scores don't appear directly on the Amazon Macie console. However, they're available in JSON representations of findings on the Macie console, the Amazon Macie API, and sensitive data discovery results that correlate to sensitive data findings. Severity scores are also included in finding events that Macie publishes to Amazon EventBridge and findings that Macie publishes to AWS Security Hub.

The topics in this section indicate how Macie determines the severity of policy findings and sensitive data findings.

### Topics

- [Severity scoring for policy findings \(p. 158\)](#)

- [Severity scoring for sensitive data findings \(p. 158\)](#)

## Severity scoring for policy findings

The severity of a policy finding is based on the nature of a potential issue with the security or privacy of your Amazon S3 data. The following table lists the severity levels that Macie assigns to each type of policy finding. For a description of each type, see [Types of findings \(p. 105\)](#).

Finding type	Severity level
Policy:IAMUser/S3BlockPublicAccessDisabled	High
Policy:IAMUser/S3BucketEncryptionDisabled	Low
Policy:IAMUser/S3BucketPublic	High
Policy:IAMUser/S3BucketReplicatedExternally	High
Policy:IAMUser/S3BucketSharedExternally	High

The severity of a policy finding doesn't change based on the number of occurrences of the finding.

## Severity scoring for sensitive data findings

The severity of a sensitive data finding is based on the nature and number of occurrences of sensitive data that Macie found in an S3 object. The following topics indicate how Macie determines the severity of each type of sensitive data finding:

- [SensitiveData:S3Object/Credentials \(p. 158\)](#)
- [SensitiveData:S3Object/CustomIdentifier \(p. 159\)](#)
- [SensitiveData:S3Object/Financial \(p. 159\)](#)
- [SensitiveData:S3Object/Personal \(p. 159\)](#)
- [SensitiveData:S3Object/Multiple \(p. 161\)](#)

For detailed information about the types of data that Macie can detect and report in sensitive data findings, see [Using managed data identifiers \(p. 37\)](#) and [Building custom data identifiers \(p. 53\)](#).

### SensitiveData:S3Object/Credentials

A **SensitiveData:S3Object/Credentials** finding indicates that an S3 object contains credentials data. For this type of finding, Macie determines severity based on the type and number of occurrences of the credentials data that Macie found in the object.

The following table indicates the severity levels that Macie assigns to findings that report occurrences of credentials data in an S3 object.

Data type	1 occurrence	2–99 occurrences	100 or more occurrences
AWS secret keys	High	High	High
OpenSSH private keys	High	High	High
PGP private keys	High	High	High

Data type	1 occurrence	2–99 occurrences	100 or more occurrences
Public-Key Cryptography Standard (PKCS) private keys	High	High	High
PuTTY private keys	High	High	High

## SensitiveData:S3Object/CustomIdentifier

A **SensitiveData:S3Object/CustomIdentifier** finding indicates that an S3 object contains data that matches one or more custom data identifiers. The object might contain more than one type of sensitive data.

For this type of finding, the severity level is always **Medium**. Severity doesn't change based on any factors, such as the number of occurrences of the data that matches a custom data identifier.

## SensitiveData:S3Object/Financial

A **SensitiveData:S3Object/Financial** finding indicates that an S3 object contains financial information. For this type of finding, Macie determines severity based on the type and number of occurrences of the financial information that Macie found in the object.

The following table indicates the severity levels that Macie assigns to findings that report occurrences of financial information in an S3 object.

Data type	1 occurrence	2–99 occurrences	100 or more occurrences
Bank account number	High	High	High
Credit card expiration date	Low	Medium	High
Credit card magnetic strip data	High	High	High
Credit card number*	High	High	High
Credit card verification code	Medium	High	High

\* The severity levels are the same for credit card numbers that are and aren't in proximity of a keyword.

If a finding reports multiple types of financial information in an object, Macie determines the finding's severity by calculating the severity for each type of financial information that Macie found, determining which type produces the highest severity, and assigning that highest severity to the finding. For example, if Macie detects 10 credit card expiration dates and 10 credit card numbers in an object, Macie assigns a **High** severity level to the finding.

## SensitiveData:S3Object/Personal

A **SensitiveData:S3Object/Personal** finding indicates that an S3 object contains personal information—personal health information (PHI), personally identifiable information (PII), or a combination of the two.

For this type of finding, Macie determines severity based on the type and number of occurrences of the personal information that Macie found in the object.

The following table indicates the severity levels that Macie assigns to sensitive data findings that report occurrences of PHI in an S3 object.

Data type	1 occurrence	2–99 occurrences	100 or more occurrences
Drug Enforcement Agency (DEA) Registration Number	High	High	High
Health Insurance Claim Number (HICN)	High	High	High
Health insurance or medical identification number	High	High	High
Healthcare Common Procedure Coding System (HCPCS) code	High	High	High
National Drug Code (NDC)	High	High	High
National Provider Identifier (NPI)	High	High	High
Unique device identifier (UDI)	Low	Medium	High

The following table indicates the severity levels that Macie assigns to sensitive data findings that report occurrences of PII in an S3 object.

Data type	1 occurrence	2–99 occurrences	100 or more occurrences
Birth date	Low	Medium	High
Driver's license identification number	Low	Medium	High
Electoral roll number	High	High	High
Full name	Low	Medium	High
Global Positioning System (GPS) coordinates	Low	Medium	Medium
Mailing address	Low	Medium	High
National identification number	High	High	High

Data type	1 occurrence	2–99 occurrences	100 or more occurrences
National Insurance Number (NINO)	High	High	High
Passport number	Medium	High	High
Permanent residence number	High	High	High
Phone number	Low	Medium	High
Social Insurance Number (SIN)	High	High	High
Social Security number (SSN)	High	High	High
Taxpayer identification or reference number	High	High	High
Vehicle identification number (VIN)	Low	Low	Medium

If a finding reports multiple types of PHI, PII, or both PHI and PII in an object, Macie determines the finding's severity by calculating the severity for each data type, determining which data type produces the highest severity, and assigning that highest severity to the finding.

For example, if Macie detects 10 full names (**Medium** severity level) and 5 passport numbers (**High** severity level) in an object, Macie assigns a **High** severity level to the finding. Similarly, if Macie detects 10 full names (**Medium** severity level) and 10 health insurance identification numbers (**High** severity level) in an object, Macie assigns a **High** severity level to the finding.

## SensitiveData:S3Object/Multiple

A **SensitiveData:S3Object/Multiple** finding indicates that an S3 object contains data spanning multiple sensitive data categories—credentials, financial information, or personal information. For this type of finding, Macie determines severity by calculating the severity for each type of sensitive data that Macie found (as indicated in the preceding topics), determining which type produces the highest severity, and assigning that highest severity to the finding.

For example, if Macie detects 10 full names (**Medium** severity level) and 10 AWS secret keys (**High** severity level) in an object, Macie assigns a **High** severity level to the finding.

# Monitoring and processing Amazon Macie findings

To support integration with other applications, services, and systems, such as monitoring or event management systems, Amazon Macie automatically publishes policy and sensitive data findings to Amazon EventBridge as events. For additional support, you can configure Macie to also publish policy and sensitive data findings to AWS Security Hub.

Amazon EventBridge, formerly called Amazon CloudWatch Events, is a serverless event bus service that delivers a stream of real-time data from applications and services, and routes that data to targets such as AWS Lambda functions, Amazon Simple Notification Service topics, and Amazon Kinesis streams. With EventBridge, you can automate monitoring and processing of certain types of events, including events that Macie publishes for findings. To learn more about EventBridge, see the [Amazon EventBridge User Guide](#). To learn about using EventBridge to monitor and process findings, see [EventBridge integration](#) (p. 162).

AWS Security Hub is a security service that provides you with a comprehensive view of your security state in AWS and helps you check your environment against security industry standards and best practices. Security Hub collects security data from multiple AWS services and supported AWS Partner Network security solutions, and it helps you analyze your security trends and identify the highest priority security issues. With Security Hub, you can analyze Macie findings as part of a broader analysis of your organization's security posture. To learn more about Security Hub, see the [AWS Security Hub User Guide](#). To learn about using Security Hub to monitor and process findings, see [Security Hub integration](#) (p. 166).

When Macie creates a finding, it automatically publishes the finding to EventBridge as a new event. Depending on the publication settings that you choose for your account, Macie can also publish the finding to Security Hub. Macie publishes each new finding immediately after it finishes processing the finding. If Macie detects a subsequent occurrence of an existing policy finding, it publishes an update to the existing EventBridge event for the finding. Depending on your publication settings, Macie can also publish the update to Security Hub. Macie publishes these updates on a recurring basis, using a publication frequency that you specify in the publication settings for your account. For details about these settings and the timing with which Macie publishes findings, see [Configuring publication settings for findings](#) (p. 173).

## Topics

- [Amazon Macie integration with Amazon EventBridge](#) (p. 162)
- [Amazon Macie integration with AWS Security Hub](#) (p. 166)
- [Configuring publication settings for Amazon Macie findings](#) (p. 173)
- [Amazon EventBridge event schema for Amazon Macie findings](#) (p. 175)

## Amazon Macie integration with Amazon EventBridge

Amazon EventBridge, formerly called Amazon CloudWatch Events, is a serverless event bus service. EventBridge delivers a stream of real-time data from applications and services, and routes that data to targets such as AWS Lambda functions, Amazon Simple Notification Service (Amazon SNS) topics, and Amazon Kinesis streams. To learn more about EventBridge, see the [Amazon EventBridge User Guide](#).

With EventBridge, you can automate monitoring and processing of certain types of events. This includes events that Amazon Macie publishes automatically for new policy findings and sensitive data findings. This also includes events that Macie publishes automatically for subsequent occurrences of existing policy findings. For details about how and when Macie publishes these events, see [Configuring publication settings for findings \(p. 173\)](#).

By using EventBridge and the events that Macie publishes for findings, you can monitor and process findings in near-real time. You can then act upon findings by using other applications and services. For example, you might use EventBridge to send specific types of new findings to an AWS Lambda function. The Lambda function might then process and send the data to your security incident and event management (SIEM) system.

In addition to automated monitoring and processing, use of EventBridge enables longer-term retention of your findings data. Macie stores findings for 90 days. With EventBridge, you can send findings data to your preferred data storage platform and store the data for as long as you like.

**Note**

For long-term retention, we strongly recommend that you also configure Macie to store all of your sensitive data discovery results in an S3 bucket. To learn more, see [Storing and retaining sensitive data discovery results \(p. 96\)](#).

**Topics**

- [Using EventBridge \(p. 163\)](#)
- [Creating EventBridge rules for finding events \(p. 163\)](#)

## Using EventBridge

With EventBridge, you create rules to specify which events you want to monitor and which targets you want to perform automated actions for those events. A *target* is a destination that EventBridge sends events to.

To automate monitoring and processing tasks for findings, you can create an EventBridge rule that automatically detects Macie finding events and sends those events to another application or service for processing or other action. You can tailor the rule to send only those events that meet certain criteria. To do this, specify criteria that derive from the [EventBridge event schema for Macie findings \(p. 175\)](#).

For example, you can create a rule that sends specific types of new findings to an AWS Lambda function. The Lambda function can then perform tasks such as: process and send the data to your SIEM system; automatically apply encryption to an S3 object; or, restrict access to an object by changing the object's access control list (ACL). Or you can create a rule that automatically sends new high-severity findings to an Amazon SNS topic, which then notifies your incident response team of the finding.

In addition to invoking Lambda functions and notifying Amazon SNS topics, EventBridge supports other types of targets and actions, such as relaying events to Amazon Kinesis streams, activating AWS Step Functions state machines, and invoking the AWS Systems Manager run command. For information about supported targets, see [Amazon EventBridge targets](#) in the *Amazon EventBridge User Guide*.

## Creating EventBridge rules for finding events

The following procedures explain how to use the Amazon EventBridge console and the [AWS Command Line Interface \(AWS CLI\)](#) to create an EventBridge rule for Macie findings. The rule detects events that use the event schema and pattern for Macie findings and sends those events to an AWS Lambda function for processing.

AWS Lambda is a compute service that you can use to run code without provisioning or managing servers. You package your code and upload it to AWS Lambda as a *Lambda function*. AWS Lambda then runs the function when the function is invoked. A function can be invoked manually by you,



automatically in response to events, or in response to requests from applications or services. For information about creating and invoking Lambda functions, see the [AWS Lambda Developer Guide](#).

#### Console

This procedure explains how to use the Amazon EventBridge console to create a rule that automatically sends all Macie finding events to a Lambda function for processing. Before you create this rule, create the Lambda function that you want the rule to use as a target. When you create the rule, you'll need to specify this function as the target for the rule.

#### To create an event rule by using the console

1. Open the Amazon EventBridge console at <https://console.aws.amazon.com/events/>.
2. In the navigation pane, under **Events**, choose **Rules**.
3. In the **Rules** section, choose **Create rule**.
4. For **Name**, enter a name for the rule. Optionally enter a description of the rule in the **Description** box.
5. Under **Define pattern**, choose **Event pattern**.
6. For **Event matching pattern**, choose **Pre-defined pattern by service**.

#### Tip

You can also create a rule that uses a custom pattern to detect and act upon only a subset of Macie finding events. This subset can be based on specific fields that Macie includes in a finding event. To learn about the available fields, see [EventBridge event schema for findings \(p. 175\)](#). To learn how to create this type of rule, see [Content-based filtering with event patterns](#) in the *Amazon EventBridge User Guide*.

7. For **Service provider**, choose **AWS**.
8. For **Service name**, choose **Macie**.
9. For **Event type**, choose **Macie Finding**.
10. Under **Select event bus**, ensure that **AWS default event bus** is selected and **Enable the rule on the selected event bus** is turned on.
11. Under **Select targets**, for **Target**, choose **Lambda function**. Then, for **Function**, choose the function that you want to send the events to.
12. (Optional) For **Configure version/alias**, enter additional settings for the function.
13. (Optional) For **Configure input**, specify which event data you want to send to the function.
14. (Optional) For **Retry policy and dead-letter queue**, specify how you want to handle events that aren't delivered to the function successfully.
15. (Optional) Enter one or more tags for the rule.
16. When you finish entering settings for the rule, choose **Create**.

#### AWS CLI

This procedure explains how to use the AWS CLI to create an EventBridge rule that sends all Macie finding events to a Lambda function for processing. In the procedure, the commands are formatted for Microsoft Windows. For Linux, macOS, or Unix, replace the caret (^) line-continuation character with a backslash (\).

Before you create this rule, create the Lambda function that you want the rule to use as a target. When you create the function, note the Amazon Resource Name (ARN) of the function. You'll need to enter this ARN when you specify the target for the rule.

#### To create an event rule by using the AWS CLI

1. To create a rule that detects events for all the findings that Macie generates, use the following EventBridge [put-rule](#) command.

```
C:\> aws events put-rule ^  
--name MacieFindings ^  
--event-pattern "{\"source\":[\"aws.macie\"]}"
```

In the preceding command, replace **MacieFindings** with the name that you want for the rule.

**Tip**

You can also create a rule that uses a custom pattern to detect and act upon only a subset of Macie finding events. This subset can be based on specific fields that Macie includes in a finding event. To learn about the available fields, see [EventBridge event schema for findings \(p. 175\)](#). To learn how to create this type of rule, see [Content-based filtering with event patterns](#) in the *Amazon EventBridge User Guide*.

If the command runs successfully, EventBridge responds with the ARN of the rule. Note this ARN. You'll need to enter it in step 3.

2. To specify the Lambda function to use as a target for the rule, use the following EventBridge [put-targets](#) command.

```
C:\> aws events put-targets ^  
--rule MacieFindings ^  
--targets Id=1,Arn=arn:aws:lambda:regionalEndpoint:accountID:function:my-findings-function
```

Where **MacieFindings** is the name that you specified for the rule in step 1, and the value for the `Arn` parameter is the ARN of the function that you want the rule to use as a target.

3. To add permissions that allow the rule to invoke the target Lambda function, use the following Lambda [add-permission](#) command.

```
C:\> aws lambda add-permission ^  
--function-name my-findings-function ^  
--statement-id Sid ^  
--action lambda:InvokeFunction ^  
--principal events.amazonaws.com ^  
--source-arn arn:aws:events:regionalEndpoint:accountId:rule:MacieFindings
```

Where:

- **my-findings-function** is the name of the Lambda function that you want the rule to use as a target.
- **Sid** is a statement identifier that you define to describe the statement in the Lambda function policy.
- `source-arn` is the ARN of the EventBridge rule.

If the command runs successfully, you receive output similar to the following:

```
{  
  "Statement": "{\"Sid\":\"sid\",  
    \"Effect\":\"Allow\",  
    \"Principal\":{\"Service\":\"events.amazonaws.com\"},  
    \"Action\":\"lambda:InvokeFunction\",  
    \"Resource\":\"arn:aws:lambda:us-east-1:111122223333:function:my-findings-function\",  
    \"Condition\":  
      {\"ArnLike\":  
        {\"AWS:SourceArn\":  
          \"arn:aws:events:us-east-1:111122223333:rule/MacieFindings\"}}}"
```

```
}
```

The `Statement` value is a JSON string version of the statement that was added to the Lambda function policy.

## Amazon Macie integration with AWS Security Hub

AWS Security Hub is a service that provides you with a comprehensive view of your security posture across your AWS environment and helps you check your environment against security industry standards and best practices. It does this partly by consuming, aggregating, organizing, and prioritizing findings from multiple AWS services and supported AWS Partner Network security solutions. Security Hub helps you analyze your security trends and identify the highest priority security issues. To learn more about Security Hub, see the [AWS Security Hub User Guide](#).

Amazon Macie integration with Security Hub enables you to publish findings from Macie to Security Hub automatically. Security Hub can then include those findings in its analysis of your security posture. This means that you can use Security Hub to monitor and process policy and sensitive data findings as part of a larger, aggregated set of findings data for your AWS environment. In other words, you can analyze Macie findings while you perform a broader analysis of your organization's security posture and remediate findings as necessary. Security Hub reduces the complexity of addressing large volumes of findings from multiple providers.

In addition, Security Hub uses a standard format for all findings, including findings from Macie. Use of this format, the *AWS Security Finding Format (ASFF)*, eliminates the need for you to perform time-consuming data conversion efforts.

### Topics

- [How Macie publishes findings to Security Hub \(p. 166\)](#)
- [Examples of Macie findings in Security Hub \(p. 169\)](#)
- [Enabling and configuring Security Hub integration \(p. 173\)](#)
- [Stopping the publication of findings to Security Hub \(p. 173\)](#)

## How Macie publishes findings to Security Hub

In Security Hub, security issues are tracked as findings. Some findings come from issues that are detected by other AWS services or by supported AWS Partner Network security solutions. Security Hub also has a set of rules that it uses to detect security issues and generate findings.

Security Hub provides tools to manage findings from all of these sources. You can view and filter lists of findings and view the details of individual findings. To learn how, see [Viewing findings](#) in the *AWS Security Hub User Guide*. You can also track the status of an investigation into a finding. To learn how, see [Taking action on findings](#) in the *AWS Security Hub User Guide*.

All findings in Security Hub use a standard JSON format called the *AWS Security Finding Format (ASFF)*. The ASFF includes details about the source of an issue, the affected resources, and the current status of a finding. For more information, see [AWS Security Finding Format \(ASFF\)](#) in the *AWS Security Hub User Guide*.

Amazon Macie is one of the AWS services that publishes findings to Security Hub.

## Types of findings that Macie publishes

Depending on the publication settings that you choose for your Macie account, Macie can publish all the findings that it creates to Security Hub, both sensitive data findings and policy findings. For information about these settings and how to change them, see [Configuring publication settings for](#)

[findings \(p. 173\)](#). By default, Macie publishes only new and updated policy findings to Security Hub. Macie doesn't publish sensitive data findings to Security Hub.

## Sensitive data findings

If you configure Macie to publish [sensitive data findings \(p. 106\)](#) to Security Hub, Macie automatically publishes each sensitive data finding that it creates for your account and it does so immediately after it finishes processing the finding. Macie does this for all sensitive data findings that aren't archived automatically by a [suppression rule \(p. 152\)](#). If you're the Macie administrator for an organization, publication is also limited to findings from sensitive data discovery jobs that you ran. Only the account that creates a job can publish sensitive data findings that the job produces.

When Macie publishes sensitive data findings to Security Hub, it uses the [AWS Security Finding Format \(ASFF\)](#), which is the standard format for all findings in Security Hub. In the ASFF, the `Types` field indicates a finding's type. This field uses a taxonomy that's slightly different from the finding type taxonomy in Macie.

The following table lists the ASFF finding type for each type of sensitive data finding that Macie can create.

Macie finding type	ASFF finding type
SensitiveData:S3Object/Credentials	Sensitive Data Identifications/Passwords/ SensitiveData:S3Object-Credentials
SensitiveData:S3Object/CustomIdentifier	Sensitive Data Identifications/PII/ SensitiveData:S3Object-CustomIdentifier
SensitiveData:S3Object/Financial	Sensitive Data Identifications/Financial/ SensitiveData:S3Object-Financial
SensitiveData:S3Object/Multiple	Sensitive Data Identifications/PII/ SensitiveData:S3Object-Multiple
SensitiveData:S3Object/Personal	Sensitive Data Identifications/PII/ SensitiveData:S3Object-Personal

## Policy findings

If you configure Macie to publish [policy findings \(p. 105\)](#) to Security Hub, Macie automatically publishes each new policy finding that it creates and it does so immediately after it finishes processing the finding. If Macie detects a subsequent occurrence of an existing policy finding, it automatically publishes an update to the existing finding in Security Hub, using a publication frequency that you specify for your account.

Macie performs these tasks for all policy findings that aren't archived automatically by a [suppression rule \(p. 152\)](#). If you're the Macie administrator for an organization, publication is also limited to policy findings for S3 buckets that are owned directly by your account. Macie doesn't publish policy findings that it creates or updates for member accounts in your organization. This helps ensure that you don't have duplicate findings data in Security Hub.

As is the case for sensitive data findings, Macie uses the AWS Security Finding Format (ASFF) when it publishes new and updated policy findings to Security Hub. In the ASFF, the `Types` field uses a taxonomy that's slightly different from the finding type taxonomy in Macie.

The following table lists the ASFF finding type for each type of policy finding that Macie can create. If Macie created or updated a policy finding in Security Hub on or after January 28, 2021, the finding has one of the following values for the ASFF `Types` field in Security Hub.

Macie finding type	ASFF finding type
Policy:IAMUser/S3BlockPublicAccessDisabled	Software and Configuration Checks/AWS Security Best Practices/Policy:IAMUser-S3BlockPublicAccessDisabled
Policy:IAMUser/S3BucketEncryptionDisabled	Software and Configuration Checks/AWS Security Best Practices/Policy:IAMUser-S3BucketEncryptionDisabled
Policy:IAMUser/S3BucketPublic	Effects/Data Exposure/Policy:IAMUser-S3BucketPublic
Policy:IAMUser/S3BucketReplicatedExternally	Software and Configuration Checks/AWS Security Best Practices/Policy:IAMUser-S3BucketReplicatedExternally
Policy:IAMUser/S3BucketSharedExternally	Software and Configuration Checks/AWS Security Best Practices/Policy:IAMUser-S3BucketSharedExternally

If Macie created or last updated a policy finding before January 28, 2021, the finding has one of the following values for the `ASFF Types` field in Security Hub:

- Policy:IAMUser/S3BlockPublicAccessDisabled
- Policy:IAMUser/S3BucketEncryptionDisabled
- Policy:IAMUser/S3BucketPublic
- Policy:IAMUser/S3BucketReplicatedExternally
- Policy:IAMUser/S3BucketSharedExternally

The values in the preceding list map directly to values for the **Finding type** (`type`) field in Macie.

**Note**

As you review and process policy findings in Security Hub, note the following exceptions:

- In certain AWS Regions, Macie began using ASFF finding types for new and updated findings as early as January 25, 2021.
- If you acted upon a policy finding in Security Hub before Macie began using ASFF finding types in your AWS Region, the value for the `ASFF Types` field of the finding will be one of the Macie finding types in the preceding list. It will not be one of the ASFF finding types in the preceding table. This is true for policy findings that you acted upon using the AWS Security Hub console or the **BatchUpdateFindings** operation of the AWS Security Hub API.

## Latency for publishing findings

When Macie creates a new policy or sensitive data finding, it publishes the finding to Security Hub immediately after it finishes processing the finding.

When Macie detects a subsequent occurrence of an existing policy finding, it publishes an update to the existing Security Hub finding. The timing of the update depends on the publication frequency that you choose for your Macie account. By default, Macie publishes updates every 15 minutes. For more information, including how to change the setting for your account, see [Configuring publication settings for findings](#) (p. 173).

## Retrying publication when Security Hub is not available

If Security Hub isn't available, Macie creates a queue of findings that haven't been received by Security Hub. When the system is restored, Macie retries publication until the findings are received by Security Hub.

## Updating existing findings in Security Hub

After Macie publishes a policy finding to Security Hub, Macie updates the finding to reflect any additional occurrences of the finding or finding activity. Macie does this only for policy findings. Sensitive data findings, unlike policy findings, are all treated as new (unique) because they derive from individual sensitive data discovery jobs.

When Macie publishes an update to a policy finding, Macie updates the value for the **Updated At** (`UpdatedAt`) field of the finding. You can use this value to determine when Macie most recently detected a subsequent occurrence of the potential policy violation that produced the finding.

Macie might also update the value for the **Types** (`Types`) field of a finding if the existing value for the field isn't an [ASFF finding type](#) (p. 167). This depends on whether you've acted upon the finding in Security Hub. If you haven't acted upon the finding, Macie changes the field's value to the appropriate ASFF finding type. If you've acted upon the finding, using either the AWS Security Hub console or the **BatchUpdateFindings** operation of the AWS Security Hub API, Macie doesn't change the field's value.

## Examples of Macie findings in Security Hub

When Macie publishes findings to Security Hub, it uses the [AWS Security Finding Format \(ASFF\)](#). This is the standard format for all findings in Security Hub. The following examples use sample data to demonstrate the structure and nature of the findings data that Macie publishes to Security Hub using this format:

- [Example of a sensitive data finding](#) (p. 169)
- [Example of a policy finding](#) (p. 171)

## Example of a sensitive data finding in Security Hub

Here's an example of a sensitive data finding that Macie published to Security Hub using the ASFF.

```
{
  "SchemaVersion": "2018-10-08",
  "Id": "5be50fce24526e670df77bc00example",
  "ProductArn": "arn:aws:securityhub:us-east-1::product/aws/macie",
  "GeneratorId": "aws/macie",
  "AwsAccountId": "111122223333",
  "Types": [
    "Sensitive Data Identifications/PII/SensitiveData:S3Object-Personal"
  ],
  "CreatedAt": "2021-06-28T23:21:49.667Z",
  "UpdatedAt": "2021-06-28T23:21:49.667Z",
  "Severity": {
    "Label": "HIGH",
    "Normalized": 70
  },
  "Title": "The S3 object contains personal information.",
  "Description": "The object contains personal information such as first or last names, addresses, or identification numbers.",
  "ProductFields": {
```

```
    "JobArn": "arn:aws:macie2:us-east-1:111122223333:classification-  
job/698e99c283a255bb2c992feceexample",  
    "S3Object.Path": "DOC-EXAMPLE-BUCKET1/2021 Sourcing.tsv",  
    "S3Object.Extension": "tsv",  
    "S3Bucket.effectivePermission": "NOT_PUBLIC",  
    "S3Object.PublicAccess": "false",  
    "S3Object.Size": "14",  
    "S3Object.StorageClass": "STANDARD",  
    "S3Bucket.allowsUnencryptedObjectUploads": "TRUE",  
    "JobId": "698e99c283a255bb2c992feceexample",  
    "aws/securityhub/FindingId": "arn:aws:securityhub:us-east-1:product/aws/  
macie/5be50fce24526e670df77bc00example",  
    "aws/securityhub/ProductName": "Macie",  
    "aws/securityhub/CompanyName": "Amazon"  
  },  
  "Resources": [  
    {  
      "Type": "AwsS3Bucket",  
      "Id": "arn:aws:s3::DOC-EXAMPLE-BUCKET1",  
      "Partition": "aws",  
      "Region": "us-east-1",  
      "Details": {  
        "AwsS3Bucket": {  
          "OwnerId":  
"7009a8971cd538e11f6b6606438875e7c86c5b672f46db45460ddcd08example",  
          "OwnerName": "johndoe",  
          "CreatedAt": "2020-12-30T18:16:25.000Z",  
          "ServerSideEncryptionConfiguration": {  
            "Rules": [  
              {  
                "ApplyServerSideEncryptionByDefault": {  
                  "SSEAlgorithm": "NONE"  
                }  
              }  
            ]  
          },  
          "PublicAccessBlockConfiguration": {  
            "BlockPublicAcls": true,  
            "BlockPublicPolicy": true,  
            "IgnorePublicAcls": true,  
            "RestrictPublicBuckets": true  
          }  
        }  
      }  
    },  
    {  
      "Type": "AwsS3Object",  
      "Id": "arn:aws:s3::DOC-EXAMPLE-BUCKET1/2021 Sourcing.tsv",  
      "Partition": "aws",  
      "Region": "us-east-1",  
      "DataClassification": {  
        "DetailedResultsLocation": "s3://macie-data-discovery-results/  
AWSLogs/111122223333/Macie/us-east-1/  
698e99c283a255bb2c992feceexample/111122223333/32b8485d-4f3a-3aa1-be33-  
aa3f0example.jsonl.gz",  
        "Result": {  
          "MimeType": "text/tsv",  
          "SizeClassified": 14,  
          "AdditionalOccurrences": false,  
          "Status": {  
            "Code": "COMPLETE"  
          },  
          "SensitiveData": [  
            {  
              "Category": "PERSONAL_INFORMATION",  
              "Detections": [  
                {  
                  "FindingId": "arn:aws:securityhub:us-east-1:product/aws/  
macie/5be50fce24526e670df77bc00example",  
                  "JobId": "698e99c283a255bb2c992feceexample",  
                  "Status": "COMPLETE"  
                }  
              ]  
            }  
          ]  
        }  
      }  
    }  
  ]  
}
```

```
{
  "Count": 1,
  "Type": "USA_SOCIAL_SECURITY_NUMBER",
  "Occurrences": {
    "Cells": [
      {
        "Row": 1,
        "Column": 1,
        "ColumnName": "Other",
        "CellReference": null
      }
    ]
  }
},
{
  "TotalCount": 1
}
],
"CustomDataIdentifiers": {
  "Detections": [
    ],
    "TotalCount": 0
  }
},
{
  "Details": {
    "AwsS3Object": {
      "LastModified": "2021-03-22T18:16:46.000Z",
      "ETag": "ebelca03ee8d006d457444445example",
      "VersionId": "SlBC72z5hArgexOJifxw_IN57EXAMPLE",
      "ServerSideEncryption": "NONE"
    }
  }
},
{
  "WorkflowState": "NEW",
  "Workflow": {
    "Status": "NEW"
  },
  "RecordState": "ACTIVE",
  "FindingProviderFields": {
    "Severity": {
      "Label": "HIGH"
    },
    "Types": [
      "Sensitive Data Identifications/PII/SensitiveData:S3Object-Personal"
    ]
  }
}
}
```

## Example of a policy finding in Security Hub

Here's an example of a new policy finding that Macie published to Security Hub in the ASFF.

```
{
  "SchemaVersion": "2018-10-08",
  "Id": "36ca8ba0-caf1-4fee-875c-37760example",
  "ProductArn": "arn:aws:securityhub:us-east-1::product/aws/macie",
  "GeneratorId": "aws/macie",
  "AwsAccountId": "111122223333",
  "Types": [
    "Software and Configuration Checks/AWS Security Best Practices/Policy:IAMUser-S3BlockPublicAccessDisabled"
  ],
}
```



```
"CreatedAt": "2021-06-28T01:20:52.313Z",
"UpdatedAt": "2021-06-28T01:20:52.313Z",
"Severity": {
  "Label": "HIGH",
  "Normalized": 70
},
"Title": "Block Public Access settings are disabled for the S3 bucket",
>Description": "All Amazon S3 block public access settings are disabled for the Amazon
S3 bucket. Access to the bucket is
controlled only by access control lists (ACLs) or bucket policies.",
"ProductFields": {
  "S3Bucket.effectivePermission": "PUBLIC",
  "S3Bucket.allowsUnencryptedObjectUploads": "FALSE",
  "aws/securityhub/FindingId": "arn:aws:securityhub:us-east-1::product/aws/
macie/36ca8ba0-caf1-4fee-875c-37760example",
  "aws/securityhub/ProductName": "Macie",
  "aws/securityhub/CompanyName": "Amazon"
},
"Resources": [
  {
    "Type": "AwsS3Bucket",
    "Id": "arn:aws:s3::DOC-EXAMPLE-BUCKET2",
    "Partition": "aws",
    "Region": "us-east-1",
    "Tags": {
      "Team": "Recruiting",
      "Division": "HR"
    },
  },
  "Details": {
    "AwsS3Bucket": {
      "OwnerId":
"7009a8971cd538e11f6b6606438875e7c86c5b672f46db45460ddcd08example",
      "OwnerName": "johndoe",
      "CreatedAt": "2020-11-25T18:24:38.000Z",
      "ServerSideEncryptionConfiguration": {
        "Rules": [
          {
            "ApplyServerSideEncryptionByDefault": {
              "SSEAlgorithm": "NONE"
            }
          }
        ]
      },
    },
    "PublicAccessBlockConfiguration": {
      "BlockPublicAcls": false,
      "BlockPublicPolicy": false,
      "IgnorePublicAcls": false,
      "RestrictPublicBuckets": false
    }
  }
],
"WorkflowState": "NEW",
"Workflow": {
  "Status": "NEW"
},
"RecordState": "ACTIVE",
"FindingProviderFields": {
  "Severity": {
    "Label": "LOW"
  },
  "Types": [
    "Software and Configuration Checks/AWS Security Best Practices/Policy:IAMUser-
S3BlockPublicAccessDisabled"
  ]
}
```

```
}  
}
```

## Enabling and configuring Security Hub integration

To use Macie integration with Security Hub, you must enable Security Hub for your AWS account. For information about how to enable Security Hub, see [Setting up AWS Security Hub](#) in the *AWS Security Hub User Guide*.

When you enable both Macie and Security Hub, the integration is enabled automatically. This means that Macie automatically begins to publish new and updated policy findings to Security Hub. You don't need to take any additional steps to configure the integration.

You can optionally customize your configuration by choosing the frequency with which Macie publishes updates to policy findings in Security Hub. You can also choose to publish sensitive data findings to Security Hub in addition to policy findings. To learn how, see [Configuring publication settings for findings](#) (p. 173).

## Stopping the publication of findings to Security Hub

To stop publishing findings to Security Hub, you can change the publication settings for your Macie account. To learn how, see [Choosing publication destinations for findings](#) (p. 174). You can also do this by using the Security Hub console or the Security Hub API. To learn how, see [Disabling and enabling the flow of findings from an integration \(console\)](#) or [Disabling the flow of findings from an integration \(Security Hub API, AWS CLI\)](#) in the *AWS Security Hub User Guide*.

# Configuring publication settings for Amazon Macie findings

To support integration with other applications, services, and systems, Amazon Macie automatically publishes both policy findings and sensitive data findings to Amazon EventBridge as events (formerly called Amazon CloudWatch Events). For information about how you can use EventBridge to monitor and process findings, see [EventBridge integration](#) (p. 162).

You can configure Macie to automatically publish findings to AWS Security Hub too, using destination options that you specify in the publication settings for your account. With these options, you can configure Macie to publish only policy findings, only sensitive data findings, or both policy and sensitive data findings to Security Hub. You can also configure Macie to stop publishing any findings to Security Hub. For information about how you can use Security Hub to monitor and process findings, see [Security Hub integration](#) (p. 166).

For policy findings, the timing with which Macie publishes a finding to another AWS service depends on whether the finding is new and on the publication frequency that you specify for your account. For sensitive data findings, the timing is always immediate—Macie publishes a sensitive data finding immediately after it finishes processing the finding. Unlike policy findings, Macie treats all sensitive data findings as new (unique) because they derive from individual sensitive data discovery jobs.

Note that Macie doesn't publish policy or sensitive data findings that are archived automatically by a [suppression rule](#) (p. 152). In other words, Macie doesn't publish suppressed findings to other AWS services.

### Topics

- [Choosing publication destinations for findings](#) (p. 174)
- [Determining the publication frequency for findings](#) (p. 174)

- [Changing the publication frequency for findings \(p. 175\)](#)

## Choosing publication destinations for findings

You can configure Macie to automatically publish policy and sensitive data findings to Security Hub in addition to EventBridge. By default, Macie publishes only new and updated policy findings to Security Hub. To change or extend the default configuration, adjust the publication destination settings for your account.

When you adjust your destination settings, you choose the categories of findings that you want Macie to publish to Security Hub—only sensitive data findings, only policy findings, or both sensitive data and policy findings. You can also choose to stop publishing any category of finding to Security Hub.

If you change your destination settings, your change applies only to the current AWS Region. If you're the Macie administrator for an organization, your change applies only to your account. It doesn't apply to any associated member accounts. For more information, see [Managing multiple accounts \(p. 184\)](#).

### To choose publication destinations for findings

1. Open the Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Settings**.
3. In the **Publication of findings** section, under **Destinations**, choose from the following options:
  - **Publish policy findings to** – Select the **Security Hub** check box to automatically publish new and updated policy findings to Security Hub. To stop publishing new and updated policy findings to Security Hub, clear this check box.
  - **Publish sensitive data findings to** – Select the **Security Hub** check box to automatically publish sensitive data findings to Security Hub. To stop publishing sensitive data findings to Security Hub, clear this check box.
4. Choose **Save**.

If you chose to publish any category of finding to Security Hub, make sure that you also enable Security Hub in the current Region and configure it to accept the findings that Macie publishes. Otherwise, you won't be able to access the findings in Security Hub. To learn how to accept findings in Security Hub, see [Managing product integrations](#) in the *AWS Security Hub User Guide*.

## Determining the publication frequency for findings

In Macie, each finding has a unique identifier. Macie uses this identifier to determine when to publish a finding to another AWS service:

- **New findings** – When Macie creates a new policy or sensitive data finding, it assigns a unique identifier to the finding as part of processing the finding. Immediately after Macie finishes processing the finding, it publishes the finding as a new EventBridge event. Depending on the publication settings for your account, Macie also publishes the finding as a new finding in Security Hub.
- **Updated findings** – When Macie detects a subsequent occurrence of an existing policy finding, it updates the existing finding by adding details about the subsequent occurrence and incrementing the count of occurrences. Macie also publishes these updates to the existing EventBridge event and, depending on the publication settings for your account, the existing Security Hub finding. Macie does this only for policy findings. Sensitive data findings, unlike policy findings, are all treated as new (unique) because they derive from individual sensitive data discovery jobs.

By default, Macie publishes updated findings every 15 minutes as part of a recurring publication cycle. This means that any policy findings that are updated after the most recent publication cycle

will be held, updated again as necessary, and included in the next publication cycle (approximately 15 minutes later). You can change this schedule by choosing a different publication frequency. For example, if you configure Macie to publish updated findings every hour and a publication occurs at 12:00, then any updates that occur after 12:00 are published at 13:00.

Note that neither of these cases applies to findings that are archived automatically by a [suppression rule](#) (p. 152). Macie doesn't publish suppressed findings to other AWS services.

## Changing the publication frequency for findings

You can change the schedule that Macie uses to publish updates to existing policy findings in other AWS services. By default, Macie publishes updated findings every 15 minutes. If you change this schedule, your change applies only to the current AWS Region. If you're the Macie administrator for an organization, your change also applies to all associated member accounts in the Region. For more information, see [Managing multiple accounts](#) (p. 184).

### To change the publication frequency for updated findings

1. Open the Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, choose **Settings**.
3. In the **Publication of findings** section, under **Update frequency for policy findings**, choose how often you want Macie to publish updated policy findings to other AWS services.
4. Choose **Save**.

## Amazon EventBridge event schema for Amazon Macie findings

To support integration with other applications, services, and systems, such as monitoring or event management systems, Amazon Macie automatically publishes findings to Amazon EventBridge as events. EventBridge, formerly called Amazon CloudWatch Events, is a serverless event bus service that delivers a stream of real-time data from applications and other AWS services to targets such as AWS Lambda functions, Amazon Simple Notification Service topics, and Amazon Kinesis streams. To learn more about EventBridge and EventBridge events, see the [Amazon EventBridge User Guide](#).

### Note

If you currently use CloudWatch Events, note that EventBridge and CloudWatch Events are the same underlying service and API. However, EventBridge includes additional features that enable you to receive events from software as a service (SaaS) applications and your own applications. Because the underlying service and API are the same, the event schema for Macie findings is also the same. In addition, you can use either console or API to create rules for Macie finding events.

Macie publishes events for all new findings and subsequent occurrences of existing policy findings, except findings that you archive automatically using [suppression rules](#) (p. 152). Each event is a JSON object that conforms to the EventBridge schema for AWS events and contains a JSON representation of a finding. Because the findings data is structured as an EventBridge event, you can more easily monitor, process, and act upon findings by using other applications, services, and tools.

### Topics

- [Event schema](#) (p. 176)
- [Event example for a policy finding](#) (p. 176)
- [Event example for a sensitive data finding](#) (p. 179)

## Event schema

The following example shows the schema of an [EventBridge event](#) for a Macie finding. For a detailed list of fields that can be included in a finding event, see the [Finding object table](#) in the *Amazon Macie API Reference*. The structure and fields of a finding event map closely to the Finding object of the Amazon Macie API.

```
{
  "version": "0",
  "id": "event ID",
  "detail-type": "Macie Finding",
  "source": "aws.macie",
  "account": "Amazon Web Services account ID (string)",
  "time": "event timestamp (string)",
  "region": "AWS Region (string)",
  "resources": [
    <-- ARNs of the resources involved in the event -->
  ],
  "detail": {
    <-- Details for a policy or sensitive data finding -->
  },
  "policyDetails": null, <-- Additional details for a policy finding or "null" for a
sensitive data finding -->
  "sample": Boolean,
  "archived": Boolean
}
```

## Event example for a policy finding

The following example uses sample data to demonstrate the structure and nature of objects and fields in an EventBridge event for a policy finding.

In this example, the event reports a subsequent occurrence of an existing policy finding: default encryption was disabled for an S3 bucket. The following fields and values can help you determine that this is the case:

- The `type` field is set to `Policy:IAMUser/S3BucketEncryptionDisabled`.
- The `createdAt` and `updatedAt` fields have different values. This is one indicator that the event reports a subsequent occurrence of an existing finding. The values for these fields would be the same if the event reported a new finding.
- The `count` field is set to 2, which indicates that this is the second occurrence of the finding.
- The `category` field is set to `POLICY`.
- The value for the `classificationDetails` field is `null`, which helps differentiate this event for a policy finding from an event for a sensitive data finding. For a sensitive data finding, this value would be a set of objects and fields that provide information about how and what sensitive data was found.

Also note that the value for the `sample` field is `true`. This value emphasizes that this is an example event for use in the documentation.

```
{
  "version": "0",
  "id": "0948ba87-d3b8-c6d4-f2da-732a1example",
  "detail-type": "Macie Finding",
  "source": "aws.macie",
  "account": "123456789012",
  "time": "2021-04-29T23:12:15Z",
  "region": "us-east-1",
```

```
"resources": [
],
"detail": {
  "schemaVersion": "1.0",
  "id": "64b917aa-3843-014c-91d8-937ffexample",
  "accountId": "123456789012",
  "partition": "aws",
  "region": "us-east-1",
  "type": "Policy:IAMUser/S3BucketEncryptionDisabled",
  "title": "Encryption is disabled for the S3 bucket",
  "description": "Encryption is disabled for the Amazon S3 bucket. The data in the
bucket isn't encrypted
using server-side encryption.",
  "severity": {
    "score": 1,
    "description": "Low"
  },
  "createdAt": "2021-04-29T15:46:02Z",
  "updatedAt": "2021-04-29T23:12:15Z",
  "count": 2,
  "resourcesAffected": {
    "s3Bucket": {
      "arn": "arn:aws:s3::DOC-EXAMPLE-BUCKET1",
      "name": "DOC-EXAMPLE-BUCKET1",
      "createdAt": "2020-04-03T20:46:56.000Z",
      "owner": {
        "displayName": "johndoe",
        "id":
"7009a8971cd538e11f6b6606438875e7c86c5b672f46db45460ddcd08example"
      },
      "tags": [
        {
          "key": "Division",
          "value": "HR"
        },
        {
          "key": "Team",
          "value": "Recruiting"
        }
      ],
      "defaultServerSideEncryption": {
        "encryptionType": "NONE",
        "kmsMasterKeyId": null
      },
      "publicAccess": {
        "permissionConfiguration": {
          "bucketLevelPermissions": {
            "accessControlList": {
              "allowsPublicReadAccess": false,
              "allowsPublicWriteAccess": false
            },
            "bucketPolicy": {
              "allowsPublicReadAccess": false,
              "allowsPublicWriteAccess": false
            },
            "blockPublicAccess": {
              "ignorePublicAcls": true,
              "restrictPublicBuckets": true,
              "blockPublicAcls": true,
              "blockPublicPolicy": true
            }
          },
          "accountLevelPermissions": {
            "blockPublicAccess": {
              "ignorePublicAcls": false,
```

```
        "restrictPublicBuckets": false,
        "blockPublicAcls": false,
        "blockPublicPolicy": false
      }
    },
    "effectivePermission": "NOT_PUBLIC"
  },
  "allowsUnencryptedObjectUploads": "FALSE"
},
"s3Object": null
},
"category": "POLICY",
"classificationDetails": null,
"policyDetails": {
  "action": {
    "actionType": "AWS_API_CALL",
    "apiCallDetails": {
      "api": "DeleteBucketEncryption",
      "apiServiceName": "s3.amazonaws.com",
      "firstSeen": "2021-04-29T15:46:02.401Z",
      "lastSeen": "2021-04-29T23:12:15.401Z"
    }
  },
  "actor": {
    "userIdentity": {
      "type": "AssumedRole",
      "assumedRole": {
        "principalId": "AROA1234567890EXAMPLE:AssumedRoleSessionName",
        "arn": "arn:aws:sts::123456789012:assumed-role/RoleToBeAssumed/MySessionName",
        "accountId": "111122223333",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE",
        "sessionContext": {
          "attributes": {
            "mfaAuthenticated": false,
            "creationDate": "2021-04-29T10:25:43.511Z"
          }
        },
        "sessionIssuer": {
          "type": "Role",
          "principalId": "AROA1234567890EXAMPLE",
          "arn": "arn:aws:iam::123456789012:role/RoleToBeAssumed",
          "accountId": "123456789012",
          "userName": "RoleToBeAssumed"
        }
      }
    }
  },
  "root": null,
  "iamUser": null,
  "federatedUser": null,
  "awsAccount": null,
  "awsService": null
},
"ipAddressDetails": {
  "ipAddressV4": "192.0.2.0",
  "ipOwner": {
    "asn": "-1",
    "asnOrg": "ExampleFindingASNOrg",
    "isp": "ExampleFindingISP",
    "org": "ExampleFindingORG"
  },
  "ipCountry": {
    "code": "US",
    "name": "United States"
  },
  "ipCity": {
```

```
        "name": "Ashburn"
      },
      "ipGeoLocation": {
        "lat": 39.0481,
        "lon": -77.4728
      }
    },
    "domainDetails": null
  }
},
"sample": true,
"archived": false
}
}
```

## Event example for a sensitive data finding

The following example uses sample data to demonstrate the structure and nature of objects and fields in an EventBridge event for a sensitive data finding.

In this example, the event reports a new sensitive data finding: an S3 object contains more than one category of sensitive data. The following fields and values can help you determine that this is the case:

- The `type` field is set to `SensitiveData:S3Object/Multiple`.
- The `createdAt` and `updatedAt` fields have the same values. Unlike policy findings, this is always the case for sensitive data findings. All sensitive data findings are considered new because they derive from individual jobs.
- The `count` field is set to 1, which indicates that this is a new finding. Unlike policy findings, this is always the case for sensitive data findings. All sensitive data findings are considered unique because they derive from individual jobs.
- The `category` field is set to `CLASSIFICATION`.
- The presence of the `jobArn` and `jobId` fields indicates that a sensitive data discovery job produced the finding. The values for these fields indicate which job produced the finding.
- The value for the `policyDetails` field is `null`, which helps differentiate this event for a sensitive data finding from an event for a policy finding. For a policy finding, this value would be a set of objects and fields that provide information about a potential policy violation.

Also note that the value for the `sample` field is `true`. This value emphasizes that this is an example event for use in the documentation.

```
{
  "version": "0",
  "id": "14ddd0b1-7c90-b9e3-8a68-6a408example",
  "detail-type": "Macie Finding",
  "source": "aws.macie",
  "account": "123456789012",
  "time": "2021-04-29T13:19:10Z",
  "region": "us-east-1",
  "resources": [

  ],
  "detail": {
    "schemaVersion": "1.0",
    "id": "4ed45d06-c9b9-4506-ab7f-18a57example",
    "accountId": "123456789012",
    "partition": "aws",
    "region": "us-east-1",
```



```
"type": "SensitiveData:S3Object/Multiple",
"title": "The S3 object contains multiple types of sensitive information.",
"description": "The object contains more than one type of sensitive information.",
"severity": {
  "score": 3,
  "description": "High"
},
"createdAt": "2021-04-29T13:19:10Z",
"updatedAt": "2021-04-29T13:19:10Z",
"count": 1,
"resourcesAffected": {
  "s3Bucket": {
    "arn": "arn:aws:s3:::DOC-EXAMPLE-BUCKET2",
    "name": "DOC-EXAMPLE-BUCKET2",
    "createdAt": "2020-05-15T20:46:56.000Z",
    "owner": {
      "displayName": "johndoe",
      "id":
"7009a8971cd538e11f6b6606438875e7c86c5b672f46db45460ddcd08example"
    },
    "tags": [
      {
        "key": "Division",
        "value": "HR"
      },
      {
        "key": "Team",
        "value": "Recruiting"
      }
    ],
    "defaultServerSideEncryption": {
      "encryptionType": "aws:kms",
      "kmsMasterKeyId": "arn:aws:kms:us-
east-1:123456789012:key/1234abcd-12ab-34cd-56ef-12345example"
    },
    "publicAccess": {
      "permissionConfiguration": {
        "bucketLevelPermissions": {
          "accessControlList": {
            "allowsPublicReadAccess": false,
            "allowsPublicWriteAccess": false
          },
          "bucketPolicy": {
            "allowsPublicReadAccess": false,
            "allowsPublicWriteAccess": false
          },
          "blockPublicAccess": {
            "ignorePublicAcls": true,
            "restrictPublicBuckets": true,
            "blockPublicAcls": true,
            "blockPublicPolicy": true
          }
        },
        "accountLevelPermissions": {
          "blockPublicAccess": {
            "ignorePublicAcls": false,
            "restrictPublicBuckets": false,
            "blockPublicAcls": false,
            "blockPublicPolicy": false
          }
        }
      },
      "effectivePermission": "NOT_PUBLIC"
    },
    "allowsUnencryptedObjectUploads": "TRUE",
  },
}
```

```
"s3Object":{
  "bucketArn": "arn:aws:s3:::DOC-EXAMPLE-BUCKET2",
  "key": "2020 Sourcing.csv",
  "path": "DOC-EXAMPLE-BUCKET2/2020 Sourcing.csv",
  "extension": ".csv",
  "lastModified": "2020-10-09T17:08:25.000Z",
  "versionId": "",
  "serverSideEncryption": {
    "encryptionType": "aws:kms",
    "kmsMasterKeyId": "arn:aws:kms:us-
east-1:123456789012:key/1234abcd-12ab-34cd-56ef-12345example"
  },
  "size": 4750,
  "storageClass": "STANDARD",
  "tags":[
    {
      "key":"Division",
      "value":"HR"
    },
    {
      "key":"Team",
      "value":"Recruiting"
    }
  ],
  "publicAccess": false,
  "etag": "6bb7fd4fa9d36d6b8fb8882caexample"
},
"category": "CLASSIFICATION",
"classificationDetails": {
  "jobArn": "arn:aws:macie2:us-east-1:123456789012:classification-
job/3ce05dbb7ec5505def334104bexample",
  "jobId": "3ce05dbb7ec5505def334104bexample",
  "result": {
    "status": {
      "code": "COMPLETE",
      "reason": null
    },
    "sizeClassified": 4750,
    "mimeType": "text/csv",
    "additionalOccurrences": true,
    "sensitiveData": [
      {
        "category": "PERSONAL_INFORMATION",
        "totalCount": 65,
        "detections": [
          {
            "type": "USA_SOCIAL_SECURITY_NUMBER",
            "count": 30,
            "occurrences": {
              "lineRanges": null,
              "offsetRanges": null,
              "pages": null,
              "records": null,
              "cells": [
                {
                  "row": 2,
                  "column": 1,
                  "columnName": "SSN",
                  "cellReference": null
                },
                {
                  "row": 3,
                  "column": 1,
                  "columnName": "SSN",
                  "cellReference": null
                }
              ]
            }
          }
        ]
      }
    ]
  }
}
```

```
    },
    {
      "row": 4,
      "column": 1,
      "columnName": "SSN",
      "cellReference": null
    }
  ]
}
},
{
  "type": "NAME",
  "count": 35,
  "occurrences": {
    "lineRanges": null,
    "offsetRanges": null,
    "pages": null,
    "records": null,
    "cells": [
      {
        "row": 2,
        "column": 3,
        "columnName": "Name",
        "cellReference": null
      },
      {
        "row": 3,
        "column": 3,
        "columnName": "Name",
        "cellReference": null
      }
    ]
  }
}
]
},
{
  "category": "FINANCIAL_INFORMATION",
  "totalCount": 30,
  "detections": [
    {
      "type": "CREDIT_CARD_NUMBER",
      "count": 30,
      "occurrences": {
        "lineRanges": null,
        "offsetRanges": null,
        "pages": null,
        "records": null,
        "cells": [
          {
            "row": 2,
            "column": 14,
            "columnName": "CCN",
            "cellReference": null
          },
          {
            "row": 3,
            "column": 14,
            "columnName": "CCN",
            "cellReference": null
          }
        ]
      }
    }
  ]
}
]
```

```
    ],
    "customDataIdentifiers": {
      "totalCount": 0,
      "detections": []
    },
    "detailedResultsLocation": "s3://macie-data-discovery-results/
AWSLogs/123456789012/Macie/us-east-1/
3ce05dbb7ec5505def334104bexample/d48bf16d-0deb-3e49-9d8c-
d407cexample.jsonl.gz"
  },
  "policyDetails": null,
  "sample": true,
  "archived": false
}
```

# Managing multiple accounts in Amazon Macie

To manage multiple Amazon Macie accounts, you choose a single account to be an administrator account for Macie. You can then associate other Macie accounts with the Macie administrator account as member accounts. There are two ways to associate accounts with a Macie administrator account: by using AWS Organizations or by sending membership invitations from Macie.

We recommend that you use AWS Organizations to manage multiple accounts. AWS Organizations is an account management service that enables AWS administrators to consolidate and centrally manage multiple accounts as a single organization. To learn more about this service, see the [AWS Organizations User Guide](#).

## Managing multiple accounts through AWS Organizations

If the account to be the Macie administrator account is part of an organization in AWS Organizations, you can designate that account as the organization's delegated Macie administrator account. You can then use the delegated Macie administrator account to enable Macie for other accounts in the organization and add those accounts as Macie member accounts.

If you already associated a Macie administrator account with member accounts by using invitations, you can designate that account as the delegated Macie administrator for the AWS organization. When you do, all currently associated member accounts remain members, which means you can take full advantage of the added benefits of managing your Macie accounts through AWS Organizations.

For more information, see [Managing multiple Amazon Macie accounts with AWS Organizations \(p. 186\)](#).

## Managing multiple accounts by invitation

If the accounts to associate aren't part of an organization in AWS Organizations, you can determine which account you want to be the Macie administrator account, and then use that administrator account to invite other accounts to become member accounts. When an invited account accepts an invitation, the account becomes a Macie member account that's associated with the Macie administrator account.

For more information, see [Managing multiple Amazon Macie accounts by using invitations \(p. 188\)](#).

## Understanding the relationship between administrator and member accounts

When you use Macie in a multiple-account environment, the Macie administrator account has access to certain metadata, S3 bucket configuration data, and policy findings for member accounts. The

administrator account can also create sensitive data discovery jobs that analyze bucket objects on behalf of member accounts.

A Macie administrator account can primarily perform the following tasks:

- Add and remove member accounts. The process by which this is done differs based on whether the accounts are associated through AWS Organizations or by invitation.
- Manage the status of Macie for associated member accounts, including enabling and suspending Macie.
- Create sensitive data discovery jobs for buckets that are owned by member accounts. Note that only the account that creates a job can access information about the job and any sensitive data findings that the job produces.

The following table provides details about the relationship between Macie administrator and member accounts. "Self" indicates that the account can't perform the action for any associated accounts. "Any" indicates that the account can perform the action for any associated accounts. "All" indicates actions that are applied to all associated accounts when they are performed by the designated account.

Action	Designation		
	Administrator	Administrator	Member
	Through AWS Organizations	By invitation	
View accounts in your organization	Any	Any	–
Enable Macie	Any	Self	Self
View policy findings	Any	Any	Self
Create sensitive data discovery jobs	Any	Any	Self
View the details of sensitive data discovery jobs <sup>1</sup>	Self	Self	Self
View sensitive data findings <sup>2</sup>	Self	Self	Self
Suppress findings	Self	Self	Self
Generate sample findings	Self	Self	Self
Set the publication frequency for findings	All	All	Self
Configure publication destinations for findings	Self	Self	Self
Configure a repository for sensitive data discovery results	Self	Self	Self
Suspend Macie <sup>3</sup>	Any	Any	Self

1. Only the account that creates a job can access information about the job. This includes job-related details in the S3 bucket inventory.
2. Only the account that creates a job can access or publish sensitive data findings that the job produces.
3. To take this action for a Macie administrator account, you must first disassociate the account from all of its member accounts in Macie.

# Managing multiple Amazon Macie accounts with AWS Organizations

When you use Amazon Macie with AWS Organizations, you designate an account as the delegated Macie administrator for the organization. Only the management account for an AWS organization can designate a delegated Macie administrator for their organization.

When an account is designated as a delegated Macie administrator, the account becomes a Macie administrator account, has Macie automatically enabled in the designated AWS Region, and is granted permission to enable and manage Macie for all the accounts that are in the organization in that Region. Additional accounts in the organization can be viewed and added as Macie member accounts that are associated with the administrator account.

If you already used invitations to set up a Macie administrator account with associated member accounts, and the member accounts are in the same AWS organization, their type changes from **By Invitation** to **Via Organizations** when you designate the existing Macie administrator account as the delegated Macie administrator for an AWS organization. If the existing member accounts aren't in the same AWS organization, their type continues to be **By Invitation**. In both cases, these previously added accounts become member accounts of the delegated Macie administrator account. You can continue to add accounts as members even if they aren't in your organization. You can do this by [sending invitations from Macie \(p. 188\)](#).

## Considerations

- There is a limit of 5,000 member accounts for each delegated Macie administrator account. However, you might have more than 5,000 accounts in your organization. If you exceed 5,000 member accounts, you'll receive notification in Amazon CloudWatch, AWS Personal Health Dashboard, and email to the administrator account.
- Although the management account for an organization can also be the delegated Macie administrator account, we don't recommend this configuration based on AWS Security best practices and the principle of least privilege. If you prefer this configuration, enable Macie for the organization's management account in at least one Region before you designate the account as the delegated Macie administrator account. Otherwise, the administrator won't be able to manage Macie settings or resources for associated member accounts.
- If you remove the delegated Macie administrator, all associated member accounts are removed as Macie members, but Macie isn't disabled for those accounts.

The topics in this section explain how to designate a delegated Macie administrator account for an AWS organization and how to add existing organization accounts as member accounts.

## Topics

- [Designating a delegated Macie administrator for an AWS organization \(p. 186\)](#)
- [Adding existing organization accounts as members \(p. 188\)](#)

## Designating a delegated Macie administrator for an AWS organization

Before you designate a delegated Macie administrator account for your AWS organization, verify that you're allowed to perform the following AWS Organizations actions:

- `organizations:DescribeOrganization`

- `organizations:EnableAWSServiceAccess`
- `organizations:ListAWSServiceAccessForOrganization`
- `organizations:RegisterDelegatedAdministrator`

These actions allow you to: retrieve information about your organization; integrate Macie with AWS Organizations; retrieve information about the AWS services that you enabled to integrate with your organization; and, administer AWS Organizations features in Macie.

To grant these permissions, append the following statement to an existing Macie policy for your account:

```
{
  "Sid": "Permissions to designate a delegated Macie administrator",
  "Effect": "Allow",
  "Action": [
    "organizations:EnableAWSServiceAccess",
    "organizations:DescribeOrganization",
    "organizations:RegisterDelegatedAdministrator",
    "organizations:ListAWSServiceAccessForOrganization"
  ],
  "Resource": "*"
}
```

If you want to designate your AWS Organizations management account as the delegated Macie administrator for your organization, the account also needs permission to perform the following AWS Identity and Access Management (IAM) action: `CreateServiceLinkedRole`. This action allows you to enable Macie.

To grant this permission, add the following statement to the IAM policy for your AWS Organizations management account:

```
{
  "Sid": "Permissions to enable Macie",
  "Effect": "Allow",
  "Action": [
    "iam:CreateServiceLinkedRole"
  ],
  "Resource": "arn:aws:iam::111122223333:role/aws-service-role/macie.amazonaws.com/AWSServiceRoleForAmazonMacie",
  "Condition": {
    "StringLike": {
      "iam:AWSServiceName": "macie.amazonaws.com"
    }
  }
}
```

In the statement, replace **111122223333** with the account ID for your AWS account.

#### Note

If you want to use Macie in a manually enabled AWS Region, also replace the value for the Macie service principal in the `Resource` element and the `iam:AWSServiceName` condition key. The value must specify the Region code for the Region. For example, if you're using Macie in the Middle East (Bahrain) Region, which has the Region code *me-south-1*, do the following:

- For the `Resource` element, replace

```
arn:aws:iam::111122223333:role/aws-service-role/macie.amazonaws.com/AWSServiceRoleForAmazonMacie
```

with



```
arn:aws:iam::111122223333:role/aws-service-role/macie.me-  
south-1.amazonaws.com/AWSServiceRoleForAmazonMacie,
```

where **111122223333** is the account ID for your AWS account.

- For the `iam:AWSServiceName` condition key, replace `macie.amazonaws.com` with `macie.me-south-1.amazonaws.com`.

After you verify your permissions, you can designate a delegated Macie administrator account for your organization. Thereafter, you need only use your organization's management account to change or remove the delegated administrator account.

### To designate a delegated Macie administrator account

1. Log in to the AWS Management Console using your AWS Organizations management account.
2. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
3. In the navigation pane, choose **Settings**.
4. Under **Delegated administrator**, enter the 12-digit account ID for the AWS account that you want to designate as the delegated Macie administrator account.
5. Choose **Delegate**.

Repeat the preceding steps in each AWS Region where your organization uses Macie. We recommend that you designate the same delegated administrator in each Region.

## Adding existing organization accounts as members

When you add an account in an AWS organization as a Macie member account, Macie is automatically enabled for that account in the current AWS Region. To add and enable Macie for those accounts in additional Regions, you must add the accounts as Macie member accounts in each additional Region.

### To add existing accounts as members

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, under **Settings**, choose **Accounts**. The **Accounts** page lists all the accounts in the organization. To quickly find specific accounts, you can sort and filter the table.
3. Select the check box for each account that you want to add as a Macie member account.
4. On the **Actions** menu, choose **Add member**.
5. Confirm that you want to add the selected accounts as members. After you do this, the status of the selected accounts changes to **Enabled**.

Repeat the preceding steps in each Region where your organization uses Macie.

## Managing multiple Amazon Macie accounts by using invitations

To manage Amazon Macie accounts that aren't associated with your account through AWS Organizations, you can use membership invitations. If you use membership invitations, your account is designated as a Macie administrator account when another account accepts your invitation to become a Macie member account.

If your account isn't a Macie administrator account, you can accept an invitation from another account. When you accept, your account becomes a Macie member account. An account can't be a Macie administrator and member account at the same time.

If accounts are associated by invitation, they have the same general administrator-to-member relationship as accounts that are associated through AWS Organizations, as described in [Understanding the relationship between administrator and member accounts \(p. 184\)](#). However, invitation-based administrator accounts can't enable Macie for member accounts.

#### Topics

- [Adding a member account \(p. 189\)](#)
- [Inviting an account \(p. 190\)](#)
- [Accepting an invitation \(p. 190\)](#)

## Adding a member account

Before you can send a Macie membership invitation to an AWS account, you must add the account. You can add accounts one at a time or in bulk.

#### To add one member account at a time

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, under **Settings**, choose **Accounts**.
3. Choose **Add accounts**.
4. For **Account ID**, enter the 12-digit account ID for the AWS account to add.
5. For **Email address**, enter the email address that's associated with the account to add.
6. Choose **Add**. When you finish adding accounts, choose **Next**.

Alternatively, you can use a CSV file to add multiple accounts at one time.

#### To add multiple member accounts in bulk

1. Using a text editor, create a CSV file as follows:
  - a. Add the following header as the first line of the CSV file:

```
Account ID,Email
```

- b. For each account, create a new line that has the account ID and email address for the account, separated by a comma. For example:

```
123456789012,user@example.com
```

- c. Save the file on your computer.
2. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
  3. In the navigation pane, under **Settings**, choose **Accounts**.
  4. Choose **Add accounts**.
  5. In the **Enter accounts** section, choose **Upload List (.csv)**.
  6. Choose **Browse**, and then select the CSV file.
  7. Choose **Add accounts**, and then choose **Next**.

## Inviting an account

The initial status of an associated account is **Created**. After you invite an account, the status of the account changes to **Email verification in progress** and then to **Invited**. After the account accepts the invitation, the status changes to **Enabled**.

### To send a membership invitation to an account

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. In the navigation pane, under **Settings**, choose **Accounts**.
3. Find the account in the list, and then choose the **Invite** link that appears in the **Status** column for it.
4. When prompted, optionally enter an invitation message. Then choose **Invite**.

## Accepting an invitation

After you accept an invitation, your account becomes a Macie member account. The account that sent the invitation becomes the Macie administrator account for your account. The Macie administrator account can see that the status of your member account has changed to **Enabled**, and can now access certain resources and perform certain tasks for your Macie account, as described in [Understanding the relationship between administrator and member accounts](#) (p. 184).

### To accept a membership invitation from an account

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. If you haven't enabled Macie yet, choose **Enable Macie**. You have to enable Macie before you can accept a membership invitation.
3. In the navigation pane, under **Settings**, choose **Accounts**.
4. To accept the invitation, choose **Accept**, and then choose **Accept invitation**. Alternatively, decline the invitation by choosing **Decline invitation**.

# Logging Amazon Macie API calls using AWS CloudTrail

Amazon Macie integrates with AWS CloudTrail, which is a service that provides a record of actions that were taken in Macie by a user, a role, or another AWS service. CloudTrail captures all API calls for Macie as events. The calls captured include calls from the Macie console and code calls to Macie API operations.

If you create a trail, you can enable continuous delivery of CloudTrail events to an Amazon Simple Storage Service (Amazon S3) bucket, including events for Macie. If you don't configure a trail, you can still view the most recent events by using **Event history** on the CloudTrail console. Using the information collected by CloudTrail, you can determine the request that was made to Macie, the IP address from which the request was made, who made the request, when it was made, and additional details.

To learn more about CloudTrail, see the [AWS CloudTrail User Guide](#).

## Macie information in CloudTrail

CloudTrail is enabled for your Amazon Web Services account when you create the account. When activity occurs in Macie, that activity is recorded in a CloudTrail event along with other Amazon Web Services events in **Event history**. You can view, search, and download recent events in your Amazon Web Services account. For more information, see [Viewing events with CloudTrail Event History](#).

For an ongoing record of events in your Amazon Web Services account, including events for Macie, create a trail. A *trail* enables CloudTrail to deliver log files to an S3 bucket. By default, when you create a trail in the console, the trail applies to all AWS Regions. The trail logs events from all Regions in the AWS partition and delivers the log files to the S3 bucket that you specify. Additionally, you can configure other AWS services to further analyze and act upon the event data collected in CloudTrail logs. For more information, see the following topics in the *AWS CloudTrail User Guide*:

- [Overview for creating a trail](#)
- [CloudTrail supported services and integrations](#)
- [Configuring Amazon SNS notifications for CloudTrail](#)
- [Receiving CloudTrail log files from multiple Regions](#)
- [Receiving CloudTrail log files from multiple accounts](#)

All Macie actions are logged by CloudTrail and are documented in the [Amazon Macie API Reference](#). For example, calls to the `ListFindings` and `CreateFindingsFilter` actions generate entries in the CloudTrail log files.

Every event or log entry contains information about who generated the request. The identity information helps you determine the following:

- Whether the request was made with root or AWS Identity and Access Management (IAM) user credentials.
- Whether the request was made with temporary security credentials for a role or federated user.
- Whether the request was made by another AWS service.

For more information, see the [CloudTrail `userIdentity` element](#).

## Understanding Macie log file entries

A trail is a configuration that enables delivery of events as log files to an S3 bucket that you specify. CloudTrail log files contain one or more log entries. An event represents a single request from any source and includes information about the requested action, the date and time of the action, request parameters, and so on. CloudTrail log files aren't an ordered stack trace of the public API calls, so they don't appear in any specific order.

The following example shows a CloudTrail log entry that demonstrates the `ListFindings` action.

```
{
  "eventVersion": "1.05",
  "userIdentity": {
    "type": "IAMUser",
    "principalId": "123456789012",
    "arn": "arn:aws:iam::123456789012:user/Mary",
    "accountId": "123456789012",
    "accessKeyId": "AKIAIOSFODNN7EXAMPLE",
    "userName": "Mary",
  },
  "eventTime": "2020-05-22T16:09:56Z",
  "eventSource": "macie2.amazonaws.com",
  "eventName": "ListFindings",
  "awsRegion": "us-east-1",
  "sourceIPAddress": "198.51.100.1",
  "userAgent": "console.amazonaws.com",
  "requestParameters": {
    "sortCriteria": {
      "attributeName": "updatedAt",
      "orderBy": "DESC"
    },
    "findingCriteria": {
      "criterion": {
        "archived": {
          "eq": [
            "false"
          ]
        },
        "category": {
          "eq": [
            "POLICY"
          ]
        }
      }
    },
    "maxResults": 10
  },
  "responseElements": null,
  "requestID": "d58af6be-1115-4a41-91f8-ace03example",
  "eventID": "ad97fac5-f7cf-4ff9-9cf2-d0676example",
  "eventType": "AwsApiCall",
  "recipientAccountId": "123456789012"
}
```

# Forecasting and monitoring Amazon Macie costs

To help you forecast and monitor your costs for using Amazon Macie, Macie calculates and provides estimated usage costs for your account. With this data, you can determine whether to adjust your use of the service or your account quotas.

You can review your estimated usage costs on the Amazon Macie console and access them programmatically with the Amazon Macie API. If you're the Macie administrator for an organization, you can review and access both aggregated data for your organization and breakdowns of the data for accounts in your organization.

If you're currently participating in the 30-day free trial of Macie, you can use this data to estimate the cost of monitoring your Amazon Simple Storage Service (Amazon S3) data for security and access control after your free trial ends. You can also check the status of your trial.

In addition to the estimated usage costs that Macie provides, you can review and monitor your actual costs by using AWS Billing and Cost Management. AWS Billing and Cost Management provides features that are designed to help you track and analyze your costs for AWS services, and manage budgets for your account or organization. It also provides features that can help you forecast usage costs based on historical data. To learn more, see the [AWS Billing and Cost Management User Guide](#).

## Topics

- [Understanding how estimated usage costs are calculated \(p. 193\)](#)
- [Reviewing estimated usage costs \(p. 194\)](#)
- [Participating in the free trial \(p. 198\)](#)

## Understanding how estimated usage costs are calculated

Amazon Macie pricing is based on two dimensions, preventative control monitoring and sensitive data discovery jobs.

### Preventative control monitoring

These costs derive from maintaining your S3 bucket inventory and evaluating and monitoring your buckets for security and access control. You're charged based on the total number of buckets that Macie can access for your account. The charges are prorated per day.

### Sensitive data discovery jobs

These costs derive from running sensitive data discovery jobs to analyze S3 objects and report sensitive data in those objects. You're charged based on the amount of uncompressed data that Macie analyzes in objects when you run a job. There's no charge for objects that Macie can't analyze for reasons such as use of an unsupported Amazon S3 storage class, use of an unsupported file or storage format, or permissions settings. For more information, see [Discovering sensitive data \(p. 36\)](#).

Note that these costs are restricted by the monthly [sensitive data discovery quota \(p. 210\)](#) for your account. (The default quota is 5 TB of data.) If a job is running and the job's analysis of eligible objects reaches this quota, Macie automatically pauses the job until you increase the quota or the next calendar month starts. If you're a Macie administrator and you run a job to analyze data for a member account, Macie automatically pauses the job until the quota is increased for the member account or the next calendar month starts.

For detailed information and examples of usage costs, see [Amazon Macie pricing](#).

When you use Macie to review your estimated usage costs, it's important to understand how the cost estimates are calculated. Consider the following:

- The estimates are reported in US Dollars and are for the current AWS Region only. If you use Macie in multiple Regions, the data isn't aggregated for all the Regions in which you use Macie.
- On the console, the estimates are inclusive for the current calendar month to date. If you query the data programmatically with the Amazon Macie API, you can choose an inclusive time range for the estimates. This can be a rolling time range of the preceding 30 days or the current calendar month to date.
- The estimates don't reflect all the discounts that might apply to your account. The exception is discounts that derive from Regional volume pricing tiers, as described in [Amazon Macie pricing](#). If your account qualifies for this type of discount, the estimates reflect that discount.

If you're the Macie administrator for an organization, the estimates don't reflect combined usage volume discounts for your organization. For information about these discounts, see [Volume discounts](#) in the *AWS Billing and Cost Management User Guide*.

- For preventative control monitoring, the estimate is based on the average daily cost for the applicable time range. The cost is prorated per day.
- For sensitive data discovery jobs, the estimate is based on the amount of uncompressed data that your jobs have analyzed thus far during the applicable time range.
- If you're the Macie administrator for an organization and you run jobs that analyze data for a member account, the estimated cost of those jobs is included in the estimate for the applicable member account. The estimated cost isn't included in the estimate for your administrator account.
- If your account is a member account in an organization and your Macie administrator runs jobs that analyze your data, the estimated cost of those jobs is included in the estimate for your account.
- The estimates don't include costs that you incur for using other AWS services with certain Macie features. For example, using customer managed AWS KMS keys to decrypt S3 objects that you want to inspect for sensitive data.

Also note that Macie provides a monthly free tier for sensitive data discovery jobs. Each month, there's no charge for you to analyze up to 1 GB of data to discover and report sensitive data in S3 objects. If you analyze more than 1 GB of data during a given month, sensitive data discovery charges begin to accrue for your account after the first 1 GB of data. If you analyze less than 1 GB of data during a given month, the remaining allocation doesn't roll over to the next month. If your account is part of an organization, the free tier applies to each individual account in your organization. In other words, there's no charge for each account in your organization to analyze up to 1 GB of data each month.

## Reviewing estimated usage costs

To review your current estimated usage costs, you can use the Amazon Macie console or the Amazon Macie API. Both the console and the API provide estimated costs for Macie pricing dimensions:

- **Preventative control monitoring** – This is the estimated cost of maintaining your S3 bucket inventory and evaluating and monitoring the buckets for security and access control.

- **Data discovery jobs** – This is the estimated cost of the sensitive data discovery jobs that you ran.

The data is reported in US Dollars and applies only to the current AWS Region. If you use the console to review the data, the cost estimates are for the current calendar month to date (inclusively). If you query the data programmatically with the Amazon Macie API, you can specify an inclusive time range for the estimates, either a rolling time range of the preceding 30 days or the current calendar month to date.

## Reviewing estimated usage costs on the Amazon Macie console

Follow these steps to review your estimated costs by using the Amazon Macie console.

### To review your estimated usage costs on the console

1. Open the Macie console at <https://console.aws.amazon.com/macie/>.
2. By using the AWS Region selector in the upper-right corner of the page, select the Region in which you want to review your estimated costs.
3. In the navigation pane, choose **Usage**.

If you have a standalone account or your account is a member account in an organization, the **Usage** page displays a breakdown of the estimated usage costs for your account.

If you're the Macie administrator for an organization, the **Usage** page lists accounts in your organization:

- In the table, the **Total** field indicates the total estimated cost for each account.
- The **Estimated costs** section shows the total estimated cost for your organization and a breakdown of those costs by pricing dimension.

To review the breakdown of estimated costs for a specific account in your organization, choose the account in the table. The **Estimated costs** section then shows this breakdown. To show this data for another account, choose the account in the table. To clear your account selection, choose **X** next to the account ID.

## Querying estimated usage costs programmatically with the Amazon Macie API

To query your estimated usage costs programmatically, you can use the following operations of the Amazon Macie API:

- **GetUsageTotals** – This operation returns total estimated usage costs for your account, grouped by usage metric. If you're the Macie administrator for an organization, this operation returns aggregated cost estimates for all the accounts in your organization. To learn more about this operation, see [Usage Totals](#) in the *Amazon Macie API Reference*.
- **GetUsageStatistics** – This operation returns usage statistics and related data for your account, grouped by account and then by usage metric. The data includes total estimated usage costs, current account quotas, and, if applicable, the date and time when the 30-day free trial started. If you're the Macie administrator for an organization, this operation returns a breakdown of the data for all the accounts in your organization. You can customize your query by sorting and filtering the query results. To learn more about this operation, see [Usage Statistics](#) in the *Amazon Macie API Reference*.

When you use either operation, you can optionally specify an inclusive time range for the data. This time range can be a rolling time range of the preceding 30 days (`PAST_30_DAYS`) or the current calendar



month to date (MONTH\_TO\_DATE). If you don't specify a time range, Macie returns the data for the preceding 30 days.

The following examples show you how to query estimated usage costs and statistics by using the [AWS Command Line Interface \(AWS CLI\)](#). You can also query the data by sending HTTPS requests directly to Macie, or by using a current version of another AWS command line tool or an AWS SDK. For information about AWS tools and SDKs, see [Tools to Build on AWS](#).

#### Examples

- [Example 1: Querying total estimated usage costs \(p. 196\)](#)
- [Example 2: Querying usage statistics \(p. 196\)](#)

## Example 1: Querying total estimated usage costs

To query total estimated usage costs by using the AWS CLI, run the [get-usage-totals](#) command and optionally specify a time range for the data. For example:

```
C:\> aws macie2 get-usage-totals --time-range MONTH_TO_DATE
```

Where MONTH\_TO\_DATE specifies the current calendar month to date as the time range for the data.

If the command runs successfully, you receive output similar to the following.

```
{
  "timeRange": "MONTH_TO_DATE",
  "usageTotals": [
    {
      "currency": "USD",
      "estimatedCost": "153.45",
      "type": "SENSITIVE_DATA_DISCOVERY"
    },
    {
      "currency": "USD",
      "estimatedCost": "10.50",
      "type": "DATA_INVENTORY_EVALUATION"
    }
  ]
}
```

Where estimatedCost is the total estimated usage cost for the associated usage metric (type): SENSITIVE\_DATA\_DISCOVERY, for analyzing S3 objects to detect sensitive data; and, DATA\_INVENTORY\_EVALUATION, for monitoring and evaluating S3 buckets for security and access control.

## Example 2: Querying usage statistics

To query usage statistics by using the AWS CLI, run the [get-usage-statistics](#) command. You can optionally sort, filter, and specify a time range for the query results. The following example retrieves usage statistics for a Macie administrator account for the preceding 30 days. The results are sorted in ascending order by account ID.

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

```
$ aws macie2 get-usage-statistics \
--sort-by '{"key":"accountId","orderBy":"ASC"}' \
```

```
--time-range PAST_30_DAYS
```

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

```
C:\> aws macie2 get-usage-statistics ^  
--sort-by={"key\":"accountId\","orderBy\":"ASC\"} ^  
--time-range PAST_30_DAYS
```

Where `PAST_30_DAYS` specifies the preceding 30 days as the time range for the data.

If the command runs successfully, Macie returns a `records` array. The array contains an object for each account that's included in the query results. For example:

```
{  
  "records": [  
    {  
      "accountId": "111122223333",  
      "freeTrialStartDate": "2020-05-20T12:26:36.917000+00:00",  
      "usage": [  
        {  
          "currency": "USD",  
          "estimatedCost": "94.53",  
          "serviceLimit": {  
            "isServiceLimited": false,  
            "unit": "TERABYTES",  
            "value": 50  
          },  
          "type": "SENSITIVE_DATA_DISCOVERY"  
        },  
        {  
          "currency": "USD",  
          "estimatedCost": "6.35",  
          "type": "DATA_INVENTORY_EVALUATION"  
        }  
      ],  
    },  
    {  
      "accountId": "444455556666",  
      "freeTrialStartDate": "2020-05-18T16:26:36.917000+00:00",  
      "usage": [  
        {  
          "currency": "USD",  
          "estimatedCost": "153.45",  
          "serviceLimit": {  
            "isServiceLimited": false,  
            "unit": "TERABYTES",  
            "value": 50  
          },  
          "type": "SENSITIVE_DATA_DISCOVERY"  
        },  
        {  
          "currency": "USD",  
          "estimatedCost": "10.50",  
          "type": "DATA_INVENTORY_EVALUATION"  
        }  
      ],  
    }  
  ],  
  "timeRange": "PAST_30_DAYS"  
}
```

Where `estimatedCost` is the total estimated usage cost for the associated usage metric (`type`) for an account: `SENSITIVE_DATA_DISCOVERY`, for analyzing S3 objects to detect sensitive data; and,

`DATA_INVENTORY_EVALUATION`, for monitoring and evaluating S3 buckets for security and access control.

## Participating in the free trial

When you enable Amazon Macie for the first time, your AWS account is automatically enrolled in the 30-day free trial of Macie. This includes individual accounts that are enabled as part of an AWS organization.

During the free trial, there's no charge for using Macie in a specific AWS Region to generate and maintain an inventory of your Amazon S3 buckets and to evaluate and monitor the buckets for security and access control. The applicable Region is the Region that's active when you enable Macie for your account. Although you can use Macie in most Regions, your account is eligible for the free trial in only one Region.

### Note

The free trial doesn't include discovery of sensitive data. This means that you'll incur charges if you create and run sensitive data discovery jobs that analyze more than 1 GB of data during the free trial. (Macie provides a monthly free tier for jobs. Each month, there's no charge for you to analyze up to 1 GB of data in S3 objects. After the first 1 GB of data, costs accrue.) You might also incur charges for other AWS services that you use with certain Macie features—for example, using customer managed AWS KMS keys to decrypt S3 objects that you want to inspect for sensitive data.

After the 30-day free trial ends, charges begin to accrue for maintaining your S3 bucket inventory and evaluating and monitoring your buckets for security and access control.

### To check your status and estimated costs during the free trial

During the free trial, you can check the status of your trial and review estimated usage costs for your account. The cost estimates are based on your use of Macie thus far during the free trial. They can help you understand what some of your usage costs might be after the free trial ends. For details about how Macie calculates these values, see [Understanding how estimated usage costs are calculated \(p. 193\)](#).

Follow these steps to review this data on the Amazon Macie console. You can also access this data programmatically by using the [GetUsageStatistics](#) operation of the Amazon Macie API.

1. Open the Macie console at <https://console.aws.amazon.com/macie/>.
2. By using the AWS Region selector in the upper-right corner of the page, select the Region in which you enrolled in the free trial.
3. In the navigation pane, choose **Usage**.

The **Usage** page indicates the number of remaining days in your free trial. It also shows a breakdown of your estimated usage costs in US Dollars:

- **Preventative control monitoring** – This is the total projected cost of maintaining your S3 bucket inventory and evaluating and monitoring your buckets for security and access control after the free trial ends.
- **Data discovery jobs** – This is the total estimated cost of any sensitive data discovery jobs that you ran. Sensitive data discovery isn't included in the free trial.

If you're the Macie administrator for an organization, the **Usage** page provides details about all the Macie accounts in your organization:

- In the table, the **Free trial** field indicates whether an account is currently participating in the free trial. (This field is empty if the free trial has ended for an account.) The **Total** field indicates the total estimated cost for each account.

- The **Estimated costs** section shows estimated costs for your organization overall.

To review the breakdown of estimated costs for a specific account in your organization, choose the account in the table. The **Estimated costs** section then shows this breakdown. To show this data for another account, choose the account in the table. To clear your account selection, choose **X** next to the account ID.

# Suspending or disabling Amazon Macie

You can suspend or disable Amazon Macie in a specific AWS Region by using the Amazon Macie console or the Amazon Macie API. Macie then stops performing all activities for your account in that Region. You aren't charged for using Macie in the Region while it's suspended or disabled.

If you suspend or disable Macie, you can re-enable it at a later time.

## Topics

- [Suspending Macie \(p. 200\)](#)
- [Disabling Macie \(p. 201\)](#)

## Suspending Macie

If you suspend Macie, it retains the session identifier, settings, and resources for your account in the applicable AWS Region. For example, your existing findings remain intact and are retained for up to 90 days. However, when you suspend Macie, it stops performing all activities for your account in the applicable Region. This includes monitoring Amazon S3 buckets and running any sensitive data discovery jobs that are currently in progress. Macie also cancels all of your sensitive data discovery jobs in the Region.

If your account is a Macie administrator account, you must disassociate your account from all of its member accounts before you suspend Macie.

After you suspend Macie, you can re-enable it. You then regain access to your settings and resources in the applicable Region, and Macie resumes its activities for your account in that Region. This includes updating the S3 bucket inventory for your account and monitoring the buckets for security and access control. This doesn't include resuming or restarting your sensitive data discovery jobs. Sensitive data discovery jobs can't be resumed or restarted after they're cancelled.

This topic explains how to suspend Macie by using the Amazon Macie console. If you prefer to do this programmatically, you can use the [Account Administration](#) resource of the Amazon Macie API.

### To suspend Macie

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. By using the AWS Region selector in the upper-right corner of the page, select the Region in which you want to suspend Macie.
3. In the navigation pane, choose **Settings**.
4. Choose **Suspend Macie**.
5. When prompted for confirmation, enter **Suspend**, and then choose **Suspend**.

To suspend Macie in multiple Regions, sign in to each additional Region, and then suspend Macie in the Region.

## Disabling Macie

When you disable Macie, Macie stops performing all activities for your account in the applicable AWS Region. This includes monitoring Amazon S3 buckets and running any sensitive data discovery jobs that are currently in progress. Macie also deletes all the existing settings and resources that it stores or maintains for your account in the applicable Region, including your findings and sensitive data discovery jobs. Resources that you stored or published to other AWS services remain intact and aren't affected—for example, sensitive data discovery results in Amazon S3 and finding events in Amazon EventBridge.

### Warning

If you disable Macie, you also permanently delete all of your existing findings, sensitive data discovery jobs, custom data identifiers, and other resources that Macie stores or maintains for your account in the applicable Region. These resources can't be recovered after they're deleted. To keep these resources and only pause your use of Macie, suspend Macie instead of disabling it.

If you want to disable Macie and your account is a Macie member account in an organization, you must disassociate your account from its Macie administrator account before you disable Macie. If your account is a Macie administrator account, you must disassociate your account from all of its member accounts before you disable Macie.

This topic explains how to disable Macie by using the Amazon Macie console. If you prefer to do this programmatically, you can use the [Account Administration](#) resource of the Amazon Macie API.

### To disable Macie

1. Open the Amazon Macie console at <https://console.aws.amazon.com/macie/>.
2. By using the AWS Region selector in the upper-right corner of the page, select the Region in which you want to disable Macie.
3. In the navigation pane, choose **Settings**.
4. Choose **Disable Macie**.
5. When prompted for confirmation, enter **Disable**, and then choose **Disable**.

To disable Macie in multiple Regions, sign in to each additional Region, and then disable Macie in the Region.

# Security in Amazon Macie

Cloud security at AWS is the highest priority. As an AWS customer, you benefit from a data center and network architecture that is built to meet the requirements of the most security-sensitive organizations.

Security is a shared responsibility between AWS and you. The [shared responsibility model](#) describes this as security of the cloud and security in the cloud:

- **Security of the cloud** – AWS is responsible for protecting the infrastructure that runs AWS services in the Amazon Web Services Cloud. AWS also provides you with services that you can use securely. Third-party auditors regularly test and verify the effectiveness of our security as part of the [AWS Compliance Programs](#). To learn about the compliance programs that apply to Macie, see [AWS Services in Scope by Compliance Program](#).
- **Security in the cloud** – Your responsibility is determined by the AWS service that you use. You are also responsible for other factors including the sensitivity of your data, your company's requirements, and applicable laws and regulations

This documentation helps you understand how to apply the shared responsibility model when using Amazon Macie. It shows you how to configure Macie to meet your security and compliance objectives. You also learn how to use other AWS services that help you to monitor and secure your Macie resources.

## Topics

- [Data protection in Amazon Macie \(p. 202\)](#)
- [Identity and access management for Amazon Macie \(p. 203\)](#)
- [Service-linked roles for Amazon Macie \(p. 205\)](#)
- [AWS managed policies for Amazon Macie \(p. 207\)](#)
- [Logging and monitoring in Amazon Macie \(p. 208\)](#)
- [Compliance validation for Amazon Macie \(p. 208\)](#)
- [Resilience in Amazon Macie \(p. 209\)](#)
- [Infrastructure security in Amazon Macie \(p. 209\)](#)

## Data protection in Amazon Macie

The AWS [shared responsibility model](#) applies to data protection in Amazon Macie. As described in this model, AWS is responsible for protecting the global infrastructure that runs all of the AWS Cloud. You are responsible for maintaining control over your content that is hosted on this infrastructure. This content includes the security configuration and management tasks for the AWS services that you use. For more information about data privacy, see the [Data Privacy FAQ](#). For information about data protection in Europe, see the [AWS Shared Responsibility Model and GDPR](#) blog post on the *AWS Security Blog*.

For data protection purposes, we recommend that you protect AWS account credentials and set up individual user accounts with AWS Identity and Access Management (IAM). That way each user is given only the permissions necessary to fulfill their job duties. We also recommend that you secure your data in the following ways:

- Use multi-factor authentication (MFA) with each account.
- Use SSL/TLS to communicate with AWS resources. We recommend TLS 1.2 or later.
- Set up API and user activity logging with AWS CloudTrail.
- Use AWS encryption solutions, along with all default security controls within AWS services.

- Use advanced managed security services such as Amazon Macie, which assists in discovering and securing personal data that is stored in Amazon S3.
- If you require FIPS 140-2 validated cryptographic modules when accessing AWS through a command line interface or an API, use a FIPS endpoint. For more information about the available FIPS endpoints, see [Federal Information Processing Standard \(FIPS\) 140-2](#).

We strongly recommend that you never put confidential or sensitive information, such as your customers' email addresses, into tags or free-form fields such as a **Name** field. This includes when you work with Macie or other AWS services using the console, API, AWS CLI, or AWS SDKs. Any data that you enter into tags or free-form fields used for names may be used for billing or diagnostic logs. If you provide a URL to an external server, we strongly recommend that you do not include credentials information in the URL to validate your request to that server.

## Encryption at rest

Amazon Macie securely stores your data at rest using AWS encryption solutions. Macie encrypts data, such as findings, using an AWS managed key from AWS Key Management Service (AWS KMS).

If you disable Macie, it permanently deletes all resources that it stores or maintains for you, such as sensitive data discovery jobs, custom data identifiers, and findings.

## Encryption in transit

Macie encrypts all data in transit between AWS services.

Amazon Macie analyzes data from Amazon S3 and exports sensitive data discovery results to an S3 bucket. After Macie gets the information that it needs from the S3 objects, they are discarded.

Macie accesses Amazon S3 using a VPC endpoint powered by AWS PrivateLink. Therefore, traffic between Macie and Amazon S3 stays on the Amazon network and does not go over the public internet. For more information, see [AWS PrivateLink](#).

# Identity and access management for Amazon Macie

AWS Identity and Access Management (IAM) is an AWS service that helps an administrator securely control access to AWS resources. IAM administrators control who can be authenticated (signed in) and authorized (have permissions) to use AWS resources. IAM enables you to create users and groups under your AWS account. You control the permissions that users have to perform tasks using AWS resources. You can use IAM for no additional charge.

By default, IAM users don't have permissions for Macie resources and operations. To allow IAM users to manage Macie resources, you must create an IAM policy that explicitly grants them permissions, and attach the policy to the IAM users or groups that require those permissions.

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 more information, see [Policies and Permissions](#) in the *IAM User Guide*.

## Policy structure

An IAM policy is a JSON document that consists of one or more statements. Each statement is structured as follows.



```
{
  "Statement": [
    {
      "Effect": "effect",
      "Action": "action",
      "Resource": "arn",
      "Condition": {
        "condition": {
          "key": "value"
        }
      }
    }
  ]
}
```

There are various elements that make up a statement:

- **Effect:** The effect can be `Allow` or `Deny`. By default, IAM users don't have permission to use resources and API actions, so all requests are denied. An explicit allow overrides the default. An explicit deny overrides any allows.
- **Action:** The action is the specific API action for which you are granting or denying permission.
- **Resource:** The resource that's affected by the action. Some API actions allow you to include specific resources in your policy that can be created or modified by the action. To specify a resource in the statement, you need to use its Amazon Resource Name (ARN).
- **Condition:** Conditions are optional. They can be used to control when your policy is in effect.

## AWS managed policies

The managed policies created by AWS grant the required permissions for common use cases. You can attach these policies to your IAM user, based on the access that they need. Each policy grants access to all or some of the API actions for Macie.

The following are the AWS managed policies for Macie:

- **AmazonMacieFullAccess** – Grants full access to Macie.
- **AmazonMacieServiceRolePolicy** – The permissions policy that's used by the [service-linked role \(p. 205\)](#) for Macie.

## API actions

In an IAM policy statement, you can specify any API action from any service that supports IAM. For Macie, use the following prefix with the name of the API action: `macie2:`. For example, `macie2:ListFindings`.

To specify multiple actions in a single statement, separate them with commas.

```
"Action": ["macie2:ListFindings", "macie2:CreateFindingsFilter"]
```

You can also specify multiple actions using wildcards. For example, you can specify all Macie API actions whose name begins with the word "Get".

```
"Action": "macie2:Get*"
```

To specify all Macie API actions, use the `*` wildcard.

```
"Action": "macie2:*"
```

For the complete list of API actions for Macie, see [Operations](#) in the *Amazon Macie API Reference*.

## Service-linked roles for Amazon Macie

Amazon Macie uses an AWS Identity and Access Management (IAM) [service-linked role](#) named *AWSServiceRoleForAmazonMacie*. The service-linked role is a unique type of IAM role that's linked directly to Macie. It's predefined by Macie and includes all the permissions that Macie needs to call other AWS services on your behalf. Macie uses the service-linked role in all the AWS Regions where Macie is available.

A service-linked role makes setting up Macie easier because you don't have to manually add the necessary permissions. Macie defines the permissions of its service-linked role, and unless defined otherwise, only Macie can assume the role. The defined permissions include the trust policy and the permissions policy, and that permissions policy can't be attached to any other IAM entity.

You can delete the Macie service-linked role only after you first disable Macie in all the Regions where it's enabled. This protects your Macie resources because you can't inadvertently remove permissions to access the resources.

## Service-linked role permissions for Macie

Macie uses the service-linked role named *AWSServiceRoleForAmazonMacie*. This role trusts the `macie.amazonaws.com` service to assume the role.

The permissions policy for the role allows Macie to perform tasks such as:

- Use Amazon S3 actions to retrieve information about S3 buckets and objects.
- Use Amazon S3 actions to retrieve S3 objects.
- Use AWS Organizations actions to describe associated accounts.
- Use Amazon CloudWatch Logs actions to log events for sensitive data discovery jobs.

The role is configured with the following permissions policy, named *AmazonMacieServiceRolePolicy*.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "cloudtrail:DescribeTrails",
        "cloudtrail:GetEventSelectors",
        "cloudtrail:GetTrailStatus",
        "cloudtrail:ListTags",
        "cloudtrail:LookupEvents",
        "iam:ListAccountAliases",
        "organizations:DescribeAccount",
        "organizations:ListAccounts",
        "s3:GetAccountPublicAccessBlock",
        "s3:ListAllMyBuckets",
        "s3:GetBucketAcl",
        "s3:GetBucketLocation",
        "s3:GetBucketLogging",
        "s3:GetBucketPolicy",
        "s3:GetBucketPolicyStatus",

```

```

        "s3:GetBucketPublicAccessBlock",
        "s3:GetBucketTagging",
        "s3:GetBucketVersioning",
        "s3:GetBucketWebsite",
        "s3:GetEncryptionConfiguration",
        "s3:GetLifecycleConfiguration",
        "s3:GetReplicationConfiguration",
        "s3:ListBucket",
        "s3:GetObject",
        "s3:GetObjectAcl",
        "s3:GetObjectTagging"
    ],
    "Resource": "*"
},
{
    "Effect": "Allow",
    "Action": [
        "cloudtrail:CreateTrail",
        "cloudtrail:StartLogging",
        "cloudtrail:StopLogging",
        "cloudtrail:UpdateTrail",
        "cloudtrail:DeleteTrail",
        "cloudtrail:PutEventSelectors"
    ],
    "Resource": "arn:aws:cloudtrail:*:*:trail/AWSMacieTrail-DO-NOT-EDIT"
},
{
    "Effect": "Allow",
    "Action": [
        "s3:CreateBucket",
        "s3:DeleteBucket",
        "s3:DeleteBucketPolicy",
        "s3:DeleteBucketWebsite",
        "s3:DeleteObject",
        "s3:DeleteObjectTagging",
        "s3:DeleteObjectVersion",
        "s3:DeleteObjectVersionTagging",
        "s3:PutBucketPolicy"
    ],
    "Resource": [
        "arn:aws:s3:::awsmacie-*",
        "arn:aws:s3:::awsmacietrail-*",
        "arn:aws:s3:::*-awsmacietrail-*"
    ]
},
{
    "Effect": "Allow",
    "Action": [
        "logs:CreateLogGroup"
    ],
    "Resource": [
        "arn:aws:logs:*:*:log-group:/aws/macie/*"
    ]
},
{
    "Effect": "Allow",
    "Action": [
        "logs:CreateLogStream",
        "logs:PutLogEvents",
        "logs:DescribeLogStreams"
    ],
    "Resource": [
        "arn:aws:logs:*:*:log-group:/aws/macie/*:log-stream:*"
    ]
}
]

```

```
}
```

For details about updates to the *AmazonMacieServiceRolePolicy* policy, see [Macie updates to AWS managed policies](#) (p. 208).

Amazon Macie and Amazon Macie Classic use the same service-linked role and permissions policy. (This helps Macie Classic users move to and use Macie.) Macie performs all the actions allowed by the policy except `CreateTrail`, `StartLogging`, `StopLogging`, `UpdateTrail`, `PutEventSelectors`, and `DeleteTrail`. Only Macie Classic performs those actions on resources, as defined by the policy.

In addition, Macie doesn't perform actions on the `arn:aws:cloudtrail:*:*:trail/AWSMacieTrail-DO-NOT-EDIT` trail or S3 buckets that have the following Amazon Resource Names: `arn:aws:s3:::awsmacie-*`, `arn:aws:s3:::awsmacietrail-*`, and `arn:aws:s3:::*-awsmacietrail-*`. Only Macie Classic performs actions on those resources, as defined by the policy.

You must configure permissions to allow an IAM entity (such as a user, group, or role) to create, edit, or delete a service-linked role. For more information, see [Service-linked role permissions](#) in the *IAM User Guide*.

## Create a service-linked role for Macie

You don't need to manually create a service-linked role for Macie. When you enable Macie, Macie automatically creates the *AWSServiceRoleForAmazonMacie* service-linked role for you.

## Edit a service-linked role for Macie

You can edit the description of the *AWSServiceRoleForAmazonMacie* service-linked role by using IAM. For more information, see [Editing a service-linked role](#) in the *IAM User Guide*.

## Delete a service-linked role for Macie

If you no longer need to use Macie, we recommend that you delete the *AWSServiceRoleForAmazonMacie* service-linked role. Before you can delete the role, you must disable Macie in each AWS Region where it's enabled. When you disable Macie, it doesn't delete the role for you. Therefore, if you enable Macie again, it can use the existing role.

You can use the IAM console, the AWS CLI, or the AWS API to manually delete the service-linked role. To do this, you must first manually clean up the resources for your service-linked role. You can then manually delete the role. For more information, see [Deleting a service-linked role](#) in the *IAM User Guide*.

If you delete this service-linked role and then need to create it again, you can use the same process to re-create the role in your account. When you enable Macie, Macie re-creates the service-linked role for you.

# AWS managed policies for Amazon Macie

To add permissions to users, groups, and roles, it is easier to use AWS managed policies than to write policies yourself. It takes time and expertise to [create IAM customer managed policies](#) that provide your team with only the permissions they need. To get started quickly, you can use our AWS managed policies. These policies cover common use cases and are available in your AWS account. For more information about AWS managed policies, see [AWS managed policies](#) in the *IAM User Guide*.

AWS services maintain and update AWS managed policies. You can't change the permissions in AWS managed policies. Services occasionally add additional permissions to an AWS managed policy to

support new features. This type of update affects all identities (users, groups, and roles) where the policy is attached. Services are most likely to update an AWS managed policy when a new feature is launched or when new operations become available. Services do not remove permissions from an AWS managed policy, so policy updates won't break your existing permissions.

Additionally, AWS supports managed policies for job functions that span multiple services. For example, the **ViewOnlyAccess** AWS managed policy provides read-only access to many AWS services and resources. When a service launches a new feature, AWS adds read-only permissions for new operations and resources. For a list and descriptions of job function policies, see [AWS managed policies for job functions](#) in the *IAM User Guide*.

## Macie updates to AWS managed policies

View details about updates to AWS managed policies for Macie since this service began tracking these changes. For automatic alerts about changes to this page, subscribe to the RSS feed on the Macie [Document history](#) (p. 213) page.

Change	Description	Date
<a href="#">AmazonMacieServiceRolePolicy</a> (p. 215) – Update to an existing policy	Macie added Amazon CloudWatch Logs actions to the <i>AmazonMacieServiceRolePolicy</i> policy. These actions allow Macie to publish log events to CloudWatch Logs for sensitive data discovery jobs.	April 13, 2021
Macie started tracking changes.	Macie started tracking changes for its AWS managed policies.	April 13, 2021

## Logging and monitoring in Amazon Macie

Amazon Macie integrates with AWS CloudTrail, which is a service that provides a record of actions that were taken in Macie by a user, a role, or another AWS service. This includes actions from the Amazon Macie console and programmatic calls to Amazon Macie API operations. By using the information collected by CloudTrail, you can determine which requests were made to Macie. For each request, you can identify when it was made, the IP address from which it was made, who made it, and additional details. For more information, see [Logging Amazon Macie API calls using AWS CloudTrail](#) (p. 191).

## Compliance validation for Amazon Macie

Third-party auditors assess the security and compliance of Amazon Macie as part of multiple AWS compliance programs. These include SOC, PCI, FedRAMP, HIPAA, and others.

For a list of AWS services in scope of specific compliance programs, see [AWS Services in Scope by Compliance Program](#). For general information, see [AWS Compliance Programs](#).

You can download third-party audit reports using AWS Artifact. For more information, see [Downloading Reports in AWS Artifact](#).

Your compliance responsibility when using Macie is determined by the sensitivity of your data, your company's compliance objectives, and applicable laws and regulations. AWS provides the following resources to help with compliance:

- [Security and Compliance Quick Start Guides](#) – These deployment guides discuss architectural considerations and provide steps for deploying security- and compliance-focused baseline environments on AWS.
- [Architecting for HIPAA Security and Compliance Whitepaper](#) – This whitepaper describes how companies can use AWS to create HIPAA-compliant applications.
- [AWS Compliance Resources](#) – This collection of workbooks and guides might apply to your industry and location.
- [Evaluating Resources with Rules](#) in the *AWS Config Developer Guide* – AWS Config assesses how well your resource configurations comply with internal practices, industry guidelines, and regulations.
- [AWS Security Hub](#) – This AWS service provides a comprehensive view of your security state within AWS that helps you check your compliance with security industry standards and best practices.

## Resilience in Amazon Macie

The AWS global infrastructure is built around AWS Regions and Availability Zones. Regions provide multiple physically separated and isolated Availability Zones, which are connected through low-latency, high-throughput, and highly redundant networking. With Availability Zones, you can design and operate applications and databases that automatically fail over between zones without interruption. Availability Zones are more highly available, fault tolerant, and scalable than traditional single or multiple data center infrastructures.

For more information about AWS Regions and Availability Zones, see [AWS Global Infrastructure](#).

## Infrastructure security in Amazon Macie

As a managed service, Macie is protected by the AWS global network security procedures that are described in the [Amazon Web Services: Overview of Security Processes](#) whitepaper.

You use AWS published API calls to access Macie through the network. Clients must support Transport Layer Security (TLS) 1.0 or later. We recommend TLS 1.2 or later. Clients must also support cipher suites with perfect forward secrecy (PFS) such as Ephemeral Diffie-Hellman (DHE) or Elliptic Curve Ephemeral Diffie-Hellman (ECDHE). Most modern systems such as Java 7 and later support these modes.

Additionally, requests must be signed using an access key ID and a secret access key that is associated with an IAM principal. Or you can use the [AWS Security Token Service](#) (AWS STS) to generate temporary security credentials to sign requests.

# Amazon Macie quotas

Your AWS account has certain default quotas, formerly referred to as *limits*, for each AWS service. These quotas are the maximum number of service resources or operations for your account. This topic lists the quotas that apply to Amazon Macie resources and operations for your account. Unless otherwise noted, each quota applies to your account in each AWS Region.

Some quotas can be increased, while others cannot. To request an increase to a quota, use the [Service Quotas console](#). To learn how to request an increase, see [Requesting a quota increase](#) in the *Service Quotas User Guide*. If a quota isn't available on the Service Quotas console, use the [service limit increase form](#) in AWS Support Center to request an increase to the quota.

## Accounts

- Member accounts by invitation: 1,000
- Member accounts through AWS Organizations: 5,000

## Findings

- Findings per run of a sensitive data discovery job: 100,000 + 5% of the Amazon S3 objects in the job

This quota applies only to the Amazon Macie console and the Amazon Macie API. There isn't a quota for the number of finding events that Macie publishes to Amazon EventBridge or the number of sensitive data discovery results that Macie creates for each run of a job.

- Detection locations per sensitive data finding: 15
- Filter and suppression rules per account: 1,000

## Sensitive data discovery

- Monthly sensitive data discovery per account: 5 TB

This quota is adjustable. To increase the quota to as much as 1,000 TB (1 PB), use the [Service Quotas console](#) to request the increase. To request an increase for more than 1 PB, use the [service limit increase form](#) to request the increase.

- Amazon S3 buckets per sensitive data discovery job: 1,000. If your account is the Macie administrator account for an organization, the buckets can span as many as 1,000 accounts in your organization.

This quota applies to a job only if you configure the job to analyze specific buckets that you select. It doesn't apply to jobs that use runtime bucket criteria to determine which buckets to analyze.

- Custom data identifiers per sensitive data discovery job: 30
- Size of an individual file to analyze:
  - Adobe Portable Document Format (.pdf) file: 1,024 MB
  - Apache Avro object container (.avro) file: 8 GB
  - Apache Parquet (.parquet) file: 8 GB
  - GNU Zip compressed archive (.gz or .gzip) file: 8 GB
  - Microsoft Excel workbook (.xls or .xlsx) file: 512 MB
  - Microsoft Word document (.doc or .docx) file: 512 MB
  - Non-binary text file: 20 GB
  - TAR archive (.tar) file: 20 GB

- ZIP compressed archive (.zip) file: 8 GB

If a file is larger than the applicable quota, Macie doesn't analyze any data in the file.

- Extraction and analysis of data in a compressed or archive file:
  - Storage size (compressed): 8 GB for a GNU Zip compressed archive (.gz or .gzip) file or ZIP compressed archive (.zip) file; 20 GB for a TAR archive (.tar) file
  - Nested archive depth: 10 levels
  - Extracted files: 1,000,000
  - Extracted bytes: 10 GB of data that uses a [supported file type or storage format \(p. 89\)](#)

If the metadata for a compressed or archive file indicates that the file contains more than 10 nested levels or exceeds the applicable quota for storage size or extracted bytes, Macie doesn't extract or analyze any data in the file.

If Macie begins to extract and analyze data in a compressed or archive file and subsequently determines that the file contains more than 1,000,000 files or exceeds the quota for extracted bytes, Macie stops analyzing data in the file and creates sensitive data findings and discovery results only for the data that was processed.

- Analysis of nested elements in structured data: 256 levels per file

This quota applies only to JSON (.json) and JSON Lines (.jsonl) files. If the nested depth of either type of file exceeds this quota, Macie doesn't analyze any data in the file.

- Detection locations in sensitive data discovery results: 1,000 per sensitive data detection type
- Detection of full names: 1,000 per file, including archive files.

After Macie detects the first 1,000 occurrences of full names in a file, Macie stops incrementing the count and reporting location data for full names.



# AWS glossary

For the latest AWS terminology, see the [AWS glossary](#) in the *AWS General Reference*.

# Document history for Amazon Macie

The following table describes the important changes to the documentation since the last release of Amazon Macie. For notification about updates to this documentation, you can subscribe to an RSS feed.

- **Latest documentation update:** October 5, 2021

update-history-change	update-history-description	update-history-date
<a href="#">New functionality (p. 213)</a>	Your <a href="#">S3 bucket inventory</a> now indicates if a bucket's permissions settings prevent Macie from retrieving information about the bucket or the bucket's objects, and evaluating and monitoring the security and privacy of the bucket's data. In addition, we updated this guide to use current terminology for AWS KMS keys and customer managed keys.	October 5, 2021
<a href="#">New functionality (p. 213)</a>	Macie now stores policy and sensitive data findings for 90 days instead of 30 days. If Macie created or updated a finding on or after August 31, 2021, you can access the finding for up to 90 days by using the Macie console or the Macie API. In certain AWS Regions, Macie began retaining findings for 90 days as early as September 27, 2021.	October 1, 2021
<a href="#">New feature (p. 213)</a>	When you <a href="#">create a sensitive data discovery job</a> , you can now specify which <a href="#">managed data identifiers</a> you want the job to use when it analyzes S3 objects. With this feature, you can tailor a job's analysis to focus on certain types of sensitive data.	September 17, 2021
<a href="#">New functionality (p. 213)</a>	Sensitive data findings now provide additional information to help you <a href="#">locate sensitive data</a> in JSON and JSON Lines files.	July 6, 2021
<a href="#">Updated functionality (p. 213)</a>	Macie now uses the <code>AwsS3Bucket</code> resource type in <a href="#">findings that it publishes to AWS Security Hub</a> . (Macie	June 28, 2021

	previously set this value to <code>AWS::S3::Bucket</code> .) <code>AwsS3Bucket</code> is the resource type value that's used for S3 buckets in the AWS Security Finding Format (ASFF).	
New feature (p. 213)	When you <a href="#">create a sensitive data discovery job</a> , you can now define <a href="#">runtime criteria</a> that determine which S3 buckets the job analyzes. With this feature, the scope of a job's analysis can dynamically adapt to changes to your bucket inventory.	May 15, 2021
New functionality (p. 213)	Your <a href="#">S3 bucket inventory</a> and the <b>Summary</b> dashboard now provide encryption metadata and statistics indicating whether buckets require server-side encryption of new objects. In addition, you can now perform on-demand refreshes of object metadata for individual buckets in your bucket inventory.	April 30, 2021
New feature (p. 213)	You can now <a href="#">use Amazon CloudWatch Logs to monitor and analyze events</a> that occur when you run sensitive data discovery jobs. To support this feature, we added CloudWatch Logs actions to the AWS managed policy for the Macie <a href="#">service-linked role</a> .	April 14, 2021
Regional availability (p. 213)	Macie is now available in the AWS Asia Pacific (Osaka) Region.	April 5, 2021
New feature (p. 213)	You can now configure Macie to <a href="#">publish sensitive data findings to AWS Security Hub</a> .	March 22, 2021
New content (p. 213)	Added information about <a href="#">monitoring and forecasting Macie costs</a> and participating in the free trial.	February 26, 2021
Updated content (p. 213)	We replaced the term <i>master account</i> with the term <i>administrator account</i> . An administrator account is used to <a href="#">centrally manage multiple accounts</a> .	February 12, 2021

New functionality (p. 213)	You can now refine the scope of sensitive data discovery jobs by <a href="#">using S3 object prefixes</a> in custom include and exclude criteria.	February 2, 2021
Updated content (p. 213)	Macie now adheres to the <a href="#">finding type taxonomy</a> of the AWS Security Finding Format (ASFF) when it publishes policy findings to AWS Security Hub.	January 28, 2021
New content (p. 213)	Added information about <a href="#">monitoring Amazon S3 data</a> and assessing the security and privacy of that data.	January 8, 2021
Regional availability (p. 213)	Macie is now available in the AWS Africa (Cape Town) Region, the AWS Europe (Milan) Region, and the AWS Middle East (Bahrain) Region.	December 21, 2020
New functionality (p. 213)	If your account is a Macie administrator account, you can now <a href="#">create and run sensitive data discovery jobs</a> that analyze data for as many as 1,000 buckets spanning as many as 1,000 accounts in your organization.	November 25, 2020
New functionality (p. 213)	Your <a href="#">S3 bucket inventory</a> now indicates whether you've configured any one-time or periodic sensitive data discovery jobs to analyze data in a bucket. If you have, it also provides details about the job that ran most recently.	November 23, 2020
New content (p. 213)	Added information about <a href="#">filtering findings</a> .	November 12, 2020
New functionality (p. 213)	Sensitive data findings now provide additional information to help you <a href="#">locate sensitive data</a> in Apache Avro object containers, Apache Parquet files, and Microsoft Excel workbooks.	November 9, 2020
New feature (p. 213)	You can now use sensitive data findings to <a href="#">locate individual occurrences of sensitive data</a> in S3 objects.	October 22, 2020
New feature (p. 213)	You can now <a href="#">pause and resume sensitive data discovery jobs</a> .	October 16, 2020

New content (p. 213)	Added details about the <a href="#">severity scoring system</a> for policy findings and sensitive data findings.	October 6, 2020
New features (p. 213)	You can now view statistics that indicate how much data Macie can analyze in individual S3 buckets when you run a sensitive data discovery job. In addition, you can now <a href="#">view the estimated cost of a job</a> when you create a job.	September 3, 2020
New content (p. 213)	Added information about <a href="#">configuring, running, and managing sensitive data discovery jobs</a> .	August 31, 2020
New functionality (p. 213)	<a href="#">Managed data identifiers</a> can now detect certain types of personally identifiable information for Brazil.	July 31, 2020
Updated content (p. 213)	Added information about the supported syntax for regular expressions in <a href="#">custom data identifiers</a> .	July 30, 2020
Updated content (p. 213)	Added keyword requirements for <a href="#">managed data identifiers</a> , and increased the <a href="#">quota</a> for the number of findings that each sensitive data discovery job can produce.	July 17, 2020
New content (p. 213)	Added information about using Amazon EventBridge and AWS Security Hub to <a href="#">monitor and process findings</a> . This includes the EventBridge event schema for findings and event examples for policy and sensitive data findings.	June 22, 2020
New content (p. 213)	Added information about <a href="#">analyzing and suppressing findings</a> .	June 17, 2020
New content (p. 213)	Added instructions for configuring Macie to <a href="#">store detailed discovery results in an S3 bucket</a> .	June 2, 2020

<a href="#">New content (p. 213)</a>	Added information about the <a href="#">types of sensitive data</a> that Macie can detect, and <a href="#">encryption requirements</a> for detecting sensitive data in Amazon S3 objects.	May 28, 2020
<a href="#">Initial release (p. 213)</a>	This is the initial release of the <i>Amazon Macie User Guide</i> .	May 13, 2020