Getting started with Amazon Macie

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

Tasks

Before you begin

When you sign up for Amazon Web Services (AWS), your account is automatically signed up for all AWS services, including Amazon Macie. However, to enable and use Macie, you first have to set up permissions that allow you to access the Amazon Macie console and API operations. You or your AWS administrator can do this by using AWS Identity and Access Management (IAM) to attach the AWS managed policy named `AmazonMacieFullAccess` to your IAM identity. To learn more, see AWS managed policies for Amazon Macie (p. 372).

Step 1: Enable Amazon Macie

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

To enable Macie

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. By using the AWS Region selector in the upper-right corner of the page, select the Region in which you want to enable and use Macie.
3. On the Amazon Macie page, choose **Get started**.
4. (Optional) When you enable Macie, Macie automatically creates a service-linked role that grants Macie the permissions that it requires to call other AWS services and monitor AWS resources on your behalf. To review the permissions policy for this role, choose **View role permissions** on the console. To learn more about this role, see Service-linked roles for Amazon Macie (p. 369).
5. Choose **Enable Macie**.

Within minutes, Macie automatically generates and begins maintaining a complete inventory of your S3 buckets in the current Region. Macie also begins evaluating and monitoring the buckets for security and access control. To learn more, see How Macie monitors Amazon S3 data security (p. 18).

Depending on your account settings, Macie also begins performing automated sensitive data discovery for your S3 buckets. Macie begins to continually identify, select, and analyze representative S3 objects in

your buckets, inspecting the objects for sensitive data. As the analyses progress, Macie provides statistics and other results that you can review, typically within 48 hours of enabling Macie for your account. You can tailor the analyses by configuring automated sensitive data discovery settings for your account. To learn more, see How automated sensitive data discovery works (p. 94).

To review aggregated statistics, choose **Summary** in the navigation pane on the console. To review details about individual S3 buckets in your inventory, choose **S3 buckets** in the navigation pane. To then display a bucket's details, choose the bucket. The details panel displays statistics and other information that provide insight into the security, privacy, and sensitivity of the bucket's data. To learn more about these details, see Reviewing your S3 bucket inventory (p. 29).

Step 2: Configure a repository for sensitive data discovery results

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

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

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

To learn how to configure this repository, see Storing and retaining sensitive data discovery results (p. 183).

Step 3: Explore sample findings

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

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

To create and review sample findings

1. Open the Amazon Macie console at https://console.aws.amazon.com/macie/.
2. In the navigation pane, choose **Settings**.
3. Under **Sample findings**, choose **Generate sample findings**. Macie generates one sample finding for each type of finding that Macie supports.

4. In the navigation pane, choose **Findings**. The **Findings** page displays findings for your account in the current AWS Region. This includes the sample findings that you created in the preceding step.
5. On the **Findings** page, locate findings whose type begins with **[SAMPLE]**.
6. To review the details of a particular sample finding, choose any field other than the check box for the finding. The details panel displays the finding's details.

To learn about each type of finding, see Types of findings (p. 197). To learn more about creating and reviewing sample findings, see Working with sample findings (p. 200).

Step 4: Create a job to discover sensitive data

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

Follow these steps to create a job that runs once, immediately after you create it, and uses default settings. To learn how to create a job that runs periodically or uses custom settings, see Creating a sensitive data discovery job (p. 147).

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 your S3 buckets in the current AWS Region.
5. Select the check box for each S3 bucket that you want the job to analyze. To find specific buckets more easily, 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**. Then choose **Next**.
9. For the **Select managed data identifiers** step, choose **Recommended**. Optionally review the table of managed data identifiers that we recommend for jobs. Then choose **Next**.

   A *managed data identifier* is a set of built-in criteria and techniques that are designed to detect a specific type of sensitive data—for example, credit card numbers, AWS secret access keys, or passport numbers for a particular country or region. To learn about managed data identifiers, see Using managed data identifiers (p. 49).
10. For the **Select custom data identifiers** step, choose **Next**.

    A *custom data identifier* is a set of criteria that you define to detect sensitive data—a regular expression (*regex*) that defines a text pattern to match and, optionally, character sequences and a proximity rule that refine the results. To learn about custom data identifiers, see Building custom data identifiers (p. 68).
11. For the **Select allow lists** step, choose **Next**.

    An *allow list* specifies text or a text pattern that you want Macie to ignore, typically sensitive data exceptions for your particular scenarios or environment. To learn about allow lists in Macie, see Defining sensitive data exceptions with allow lists (p. 74).

12. For the **Enter general settings** step, enter a name and, optionally, a description of the job. Then choose **Next**.
13. 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 about this estimate, see Forecasting the cost of a sensitive data discovery job (p. 173).
14. When you finish reviewing and verifying the job's settings, choose **Submit**.

Macie immediately starts running the job. To learn how to monitor the job, see checking the status of sensitive data discovery jobs (p. 170).

Step 5: Review your findings

Amazon Macie automatically monitors your S3 buckets for security and access control, and it creates policy findings to report potential issues with the security or privacy of your buckets. If you create and run a sensitive data discovery job or configure Macie to perform automated sensitive data discovery, Macie also creates sensitive data findings to report sensitive data that it detects in S3 objects. To learn more about findings, see Analyzing findings (p. 196).

Follow these steps to review 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**. The **Findings** page displays findings for your account in the current AWS Region.
3. (Optional) To filter the findings by specific criteria, enter the criteria in the filter bar above the table. To learn more about filters, see Filtering findings (p. 206).
4. To review the details of a particular finding, choose any field other than the check box for the finding. The details panel displays the finding's details.

To learn more, including how to group and filter findings, see Reviewing findings (p. 203).