Amazon Macie – Classification Job Configuration Guide

# 📘 Purpose

This guide outlines the steps to create a classification job in Amazon Macie for scanning sensitive data in an S3 bucket. The goal is to identify PII/PHI and route those findings securely to a discovery-results bucket using KMS encryption.

# 🛠️ Steps to Create a Macie Classification Job

1. Navigate to Amazon Macie in the AWS Console.

2. Go to the 'Jobs' tab and click 'Create job'.

3. Select the S3 bucket that stores your sensitive data (e.g., `org-sensitive-data-bucket`).

4. Choose the bucket scope (entire bucket or a specific prefix). Avoid adding a prefix unless necessary.

5. Under 'Managed data identifiers', select 'Use all recommended managed data identifiers'.

6. Choose the scan frequency: 'One-time job' for single scans or 'Scheduled' for continuous scanning.

7. Name the job (e.g., `scan-sensitive-data-test-1`) and provide a description.

8. Configure Discovery Results Export by selecting your encrypted results bucket (e.g., `macie-discovery-results-mvs`) and the correct CMK.

9. Review permissions. Ensure that Macie can access both the source bucket and the results bucket and that the CMK is properly authorized.

10. Click 'Create and run job'.

# 📌 Notes & Best Practices

- Ensure the selected files are ≥ 1 KB and in a readable format (e.g., .txt, .csv, .json).  
- Macie findings may take several minutes to appear after the job completes.  
- Confirm that bucket policies and KMS key policies allow `macie.amazonaws.com` access.  
- Use EventBridge and SNS to set up alerting based on Macie findings.