Service Name

Amazon Macie

Service Overview

Amazon Macie is a data security service that uses machine learning and pattern matching to discover, monitor, and protect sensitive data in Amazon S3. It automates the discovery and reporting of sensitive data such as personally identifiable information (PII), financial data, and credentials. Macie evaluates S3 buckets for security and access control issues, provides continuous monitoring capabilities, and generates detailed findings for potential security risks. The service integrates with AWS Security Hub and Amazon EventBridge to enable automated remediation workflows and comprehensive security posture management across organizations.

Key Use Cases

- Data privacy compliance: Automatically discover and classify sensitive data to meet regulatory requirements like GDPR, CCPA, and HIPAA
- Security posture monitoring: Continuously evaluate S3 bucket configurations for security and access control vulnerabilities
- Risk assessment and remediation: Identify unencrypted or publicly accessible buckets containing sensitive data for immediate remediation
- Multi-account data governance: Centrally manage sensitive data discovery across organizational accounts using AWS Organizations integration
- Incident response and forensics: Provide detailed location information for sensitive data to support security investigations and audit requirements

Features

- **Automated Sensitive Data Discovery**: Continuously evaluates S3 bucket inventories and analyzes representative objects for sensitive data using machine learning
- **Managed Data Identifiers**: Built-in criteria for detecting over 100 types of sensitive data including PII, financial information, and credentials
- **Custom Data Identifiers**: User-defined regex patterns to detect organization-specific sensitive data types
- **Policy Findings**: Detects changes to S3 bucket policies that reduce security or privacy posture
- **Sensitive Data Discovery Jobs**: On-demand and scheduled analysis of specific S3 buckets with customizable scope and frequency
- **Multi-Account Management**: Centralized management across AWS Organizations with delegated administrator capabilities

- **Allow Lists**: Define text patterns and exceptions that Macie should ignore during analysis
- **Integration Capabilities**: Publishes findings to EventBridge and Security Hub for automated workflows and centralized security management

Value Proposition

Amazon Macie provides unmatched automation and scale for sensitive data discovery in S3 environments, leveraging advanced machine learning to reduce manual effort and false positives. Unlike traditional data loss prevention solutions, Macie offers native AWS integration with seamless setup through a single API call or console selection:

- The service provides comprehensive coverage with over 100 built-in sensitive data types while supporting custom detection patterns for organization-specific requirements
- Macie's multi-account support through AWS Organizations enables centralized governance at enterprise scale, while its pay-as-you-analyze pricing model ensures cost efficiency
- The service's continuous monitoring capabilities and integration with Security Hub and EventBridge create automated security workflows that competitors cannot match in the AWS ecosystem

Service Integration

Amazon Macie integrates deeply with core AWS services to provide comprehensive data security management. It natively monitors Amazon S3 buckets and objects, leveraging S3 bucket policies and access controls for security evaluation:

- Macie publishes findings as events to Amazon EventBridge, enabling automated remediation through Lambda functions, SNS notifications, and Step Functions workflows
- AWS Security Hub integration provides centralized security posture management and compliance reporting across the organization
- AWS Organizations support enables multi-account deployment and management with delegated administrator capabilities
- Macie integrates with AWS CloudTrail for API call logging and audit trails, AWS IAM for access control and permissions management, and AWS KMS for analyzing encrypted objects
- The service also works with AWS Backup for data protection strategies and supports VPC endpoints for private connectivity

Onboarding and Offboarding

Customers can enable Amazon Macie with a single selection in the AWS Management Console or one API call, requiring minimal setup configuration. The service automatically

generates an inventory of S3 general purpose buckets and begins evaluating them for security and access control upon enablement:

- Users must configure IAM permissions for console and API access, then select their desired AWS region and follow on-screen instructions
- Macie immediately starts automated sensitive data discovery, analyzing representative objects in S3 buckets
- For long-term retention of discovery results beyond the default 90-day storage, customers should configure an S3 repository
- Multi-account organizations can enable trusted access between Macie and AWS Organizations, designating a delegated administrator account for centralized management across member accounts

Getting Started Guide: https://docs.aws.amazon.com/macie/latest/user/getting-started.html

Data Removal

When offboarding from Amazon Macie, customers can disable the service through the console or API, which deletes all Macie-specific settings and resources for the account. Sensitive data discovery results are automatically deleted after 90 days unless stored in a customer-managed S3 repository. Customers retain control over findings data published to EventBridge and Security Hub, which must be managed separately. To completely remove data, customers should delete any configured S3 repositories containing discovery results and remove EventBridge rules or Security Hub integrations that process Macie findings.

Backup/Restore and Disaster Recovery

Amazon Macie provides data resiliency through multiple mechanisms but does not offer traditional backup capabilities. The service stores sensitive data discovery results in customer-specified S3 buckets for long-term retention and cross-region replication. Macie publishes findings to EventBridge, allowing customers to send findings data to preferred storage platforms for extended retention. The service leverages AWS global infrastructure with multiple Availability Zones for high availability and fault tolerance. Customers can use Macie API operations to retrieve findings data programmatically and integrate with external backup systems.

Backup Capabilities

Amazon Macie does not provide traditional backup capabilities but offers data retention mechanisms. Discovery results are stored for 90 days by default and can be retained long-term in customer-managed S3 buckets:

- The service does not integrate with AWS Backup as it is not a data storage service but rather analyzes existing S3 data

- Customers must implement their own backup strategies for Macie configurations, custom data identifiers, and findings data through API exports and S3 storage

Disaster Recovery

Amazon Macie does not directly integrate with AWS Elastic Disaster Recovery as it is a security service rather than a workload requiring recovery. The service inherently provides resilience through AWS global infrastructure and multi-region availability. For disaster recovery planning, customers can replicate Macie configurations across regions using Infrastructure as Code, export findings data to multiple regions via S3 cross-region replication, and maintain EventBridge rules in multiple regions for continued security monitoring during regional outages.

Recovery Objectives

Amazon Macie supports recovery objectives by providing automated data discovery that can quickly re-establish security posture visibility after outages. The service's API-driven configuration enables rapid redeployment across regions, supporting low RTO requirements. For RPO objectives, Macie's continuous monitoring and real-time findings publication to EventBridge ensure minimal data loss during incident response scenarios. Organizations can achieve faster recovery by maintaining Macie configurations in multiple regions and implementing cross-region findings replication through S3 and EventBridge integration.

Service Constraints

Service Quotas: https://docs.aws.amazon.com/macie/latest/user/macie-quotas.html

FAQ: https://aws.amazon.com/macie/faq/

Technical Requirements

Service Documentation: https://docs.aws.amazon.com/macie/latest/user/what-is-macie.html

User Guide: https://docs.aws.amazon.com/macie/latest/user/getting-started.html

API Reference: https://docs.aws.amazon.com/macie/latest/APIReference/welcome.html

Pricing Overview

Please see the AWS UK G Cloud 15 Pricing Document accompanying this service in the Digital Marketplace.

Public Pricing: https://aws.amazon.com/macie/pricing/

Cost Optimization

Amazon Macie offers several unique cost optimization features including automated sensitive data discovery that samples representative objects rather than analyzing entire datasets, reducing analysis costs while maintaining coverage. Customers can exclude specific S3 buckets from automated discovery to focus spending on critical data stores:

- The service provides cost estimation tools during sensitive data discovery job creation, allowing customers to forecast expenses before execution
- Macie's pay-as-you-analyze pricing model with monthly quotas helps control costs, and the 30-day free trial includes up to 150GB of automated discovery analysis
- Organizations can optimize costs by configuring discovery scope, using allow lists to reduce false positives, and leveraging automated discovery over manual jobs where appropriate

Savings Considerations

Customers can achieve significant cost savings with Amazon Macie by leveraging automated discovery instead of running comprehensive manual discovery jobs, which reduces the volume of data analyzed. Strategic use of bucket exclusions and scoped discovery jobs prevents analysis of non-sensitive data repositories:

- The service's machine learning capabilities reduce operational costs by minimizing false positives and manual review efforts
- Macie's integration with existing AWS services eliminates the need for third-party data classification tools and associated licensing costs
- The monthly free tier of 1GB analysis and volume discounts for large organizations provide additional savings opportunities
- Customers should monitor usage through the Macie console's cost estimation features and optimize discovery job frequency based on data change patterns to maximize cost efficiency