AWS Security Series

Part 7
AWS Security: Data Classification and Protection with AWS Macie

Nitin Sharma
CyberSecurity and DevSecOps Engineer
LinkedIn: linkedin.com/in/nitins87
Quora: quora.com/profile/NitinS-1
Blog: 4hathacker.in

Contents

$ whoami

$ CIA Triad and Data Classification

$ Data Classification Criteria

$ “Sensitive” Personal Data: IMPORTANT

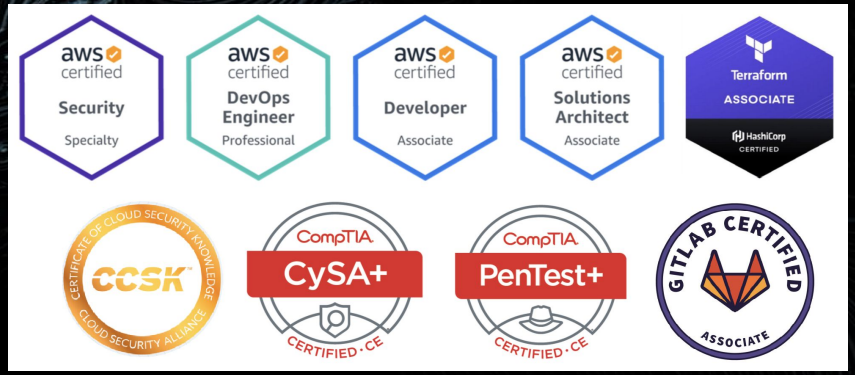
$ Data Classification: AWS Well Architected Framework

$ Amazon Macie: Intro, Features, Findings & Identifier

$ Amazon Macie Demo

$ whoami

- Cybersecurity and DevSecOps professional experienced in Cloud Security, Container Security and DevOps Research
- Certifications:



- Published author for "Securing Docker - The Attack & Defense Ways" book under CyberSecrets Publication
- Half Marathon runner, Cyclist and Fitness Enthusiast
- Helping out beginners in Cloud, DevOps and CyberSec at Quora

AWSUGNCR
aws
AWS USER GROUP NCR

$ whoami (all the way long...)

What I do everytime with teams...

What teams think about me...

$ CIA Triad and Data Classification

The CIA Triad		
What Is the CIA?		
Confidentiality	**Integrity**	**Availability**
I send you a message, and no one else knows what that message is.	I send you a message, and you receive exactly what I sent you	I send you a message, and you receive it
What's The Purpose of the CIA?		
Data is not disclosed	Data is not tampered	Data is available
How Do You Achieve the CIA?		
e.g., Encryption	e.g., Hashing, Digital signatures	e.g., Backups, redundant systems
Opposite of CIA		
Disclosure	Alteration	Destruction

(Source: Cyber News Box)

$ CIA Triad and Data Classification

	POTENTIAL IMPACT		
Security Objective	**LOW**	**MODERATE**	**HIGH**
Confidentiality Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information. [44 U.S.C., SEC. 3542]	The unauthorized disclosure of information could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized disclosure of information could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized disclosure of information could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.
Integrity Guarding against improper information modification or destruction, and includes ensuring information non-repudiation and authenticity. [44 U.S.C., SEC. 3542]	The unauthorized modification or destruction of information could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized modification or destruction of information could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized modification or destruction of information could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.
Availability Ensuring timely and reliable access to and use of information. [44 U.S.C., SEC. 3542]	The disruption of access to or use of information or an information system could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The disruption of access to or use of information or an information system could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The disruption of access to or use of information or an information system could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.

(Source: NIST FIPS-199 Potential Impact Definition)

- **Data Classification** is a fundamental step to protect proprietary information.
- **Data Sensitivity** refers to a measure of the importance assigned to data/information by its owner, for the purpose of denoting its need for protection.

$ CIA Triad and Data Classification

Government Data Classification

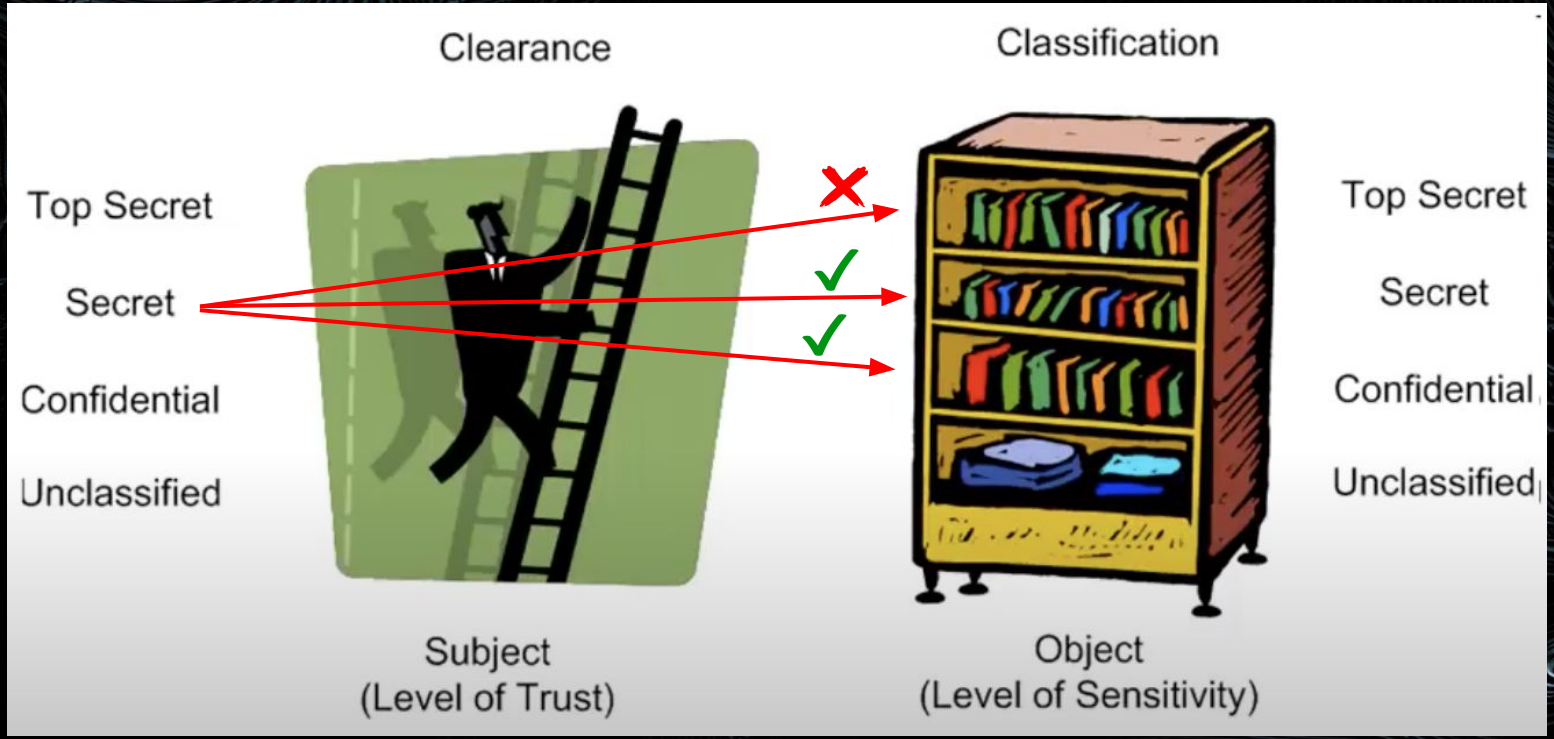
Classification	Description
Top Secret	Disclosure of top secret data would cause severe damage to national security.
Secret	Disclosure of secret data would cause serious damage to national security. This data is considered less sensitive than data classified as top secret.
Confidential	Confidential data is usually data that is exempt from disclosure under laws such as the Freedom of Information Act but is not classified as national security data.
Sensitive But Unclassified (SBU)	SBU data is data that is not considered vital to national security, but its disclosure would do some harm. Many agencies classify data they collect from citizens as SBU. In Canada, the SBU classification is referred to as protected (A, B, C).
Unclassified	Unclassified is data that has no classification or is not sensitive.

Commercial Data Classification

Classification	Description
Sensitive	Data that is to have the most limited access and requires a high degree of integrity. This is typically data that will do the most damage to the organization should it be disclosed.
Confidential	Data that might be less restrictive within the company but might cause damage if disclosed.
Private	Private data is usually compartmental data that might not do the company damage but must be keep private for other reasons. Human resources data is one example of data that can be classified as private.
Proprietary	Proprietary data is data that is disclosed outside the company on a limited basis or contains information that could reduce the company's competitive advantage, such as the technical specifications of a new product.
Public	Public data is the least sensitive data used by the company and would cause the least harm if disclosed. This could be anything from data used for marketing to the number of employees in the company.

$ CIA Triad and Data Classification

How Data Classification works ?



$ Data Classification Criteria

- Who should be able to access or maintain the data ? (Access Controls)
- Which laws, regulations, directives, or liability might be required in protecting the data ? (Compliance Controls)
- For Govt. Organizations, what would the level of damage be if the data was disclosed or corrupted ? (Risk Management and Data Laws)
- Where is the data to be stored ? (ISMS)
- What is the value of usefulness of the data ? (Requirement Constraints)

$ "Sensitive" Personal Data: IMPORTANT

1. personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs;

2. trade-union membership;

3. genetic data, biometric data processed solely to identify a human being;

4. health-related data;

5. data concerning a person's sex life or sexual orientation.

$ "Sensitive" Personal Data: IMPORTANT

Personally Identifiable Information

PDF | Kindle | RSS

Object classification by personally identifiable information (PII) is based on recognizing any personally identifiable artifacts based on industry standards such as NIST-80-122 and FIPS 199. Macie Classic can recognize the following PII artifacts:

- Full names
- Mailing addresses
- Email addresses
- Credit card numbers
- IP addresses (IPv4 and IPv6)
- Drivers license IDs (USA)
- National identification numbers (USA)
- Birth dates

As part of PII object classification, Macie Classic also assigns each matching object a PII impact of high, moderate, and low using the following criteria:

- High
  - >= 1 full name and credit card
  - >= 50 names or emails and any combination of other PII
- Moderate
  - >= 5 names or emails and any combination of other PII
- Low
  - 1–5 names or emails and any combination of PII
  - Any quantity of PII attributes above (without names or emails)

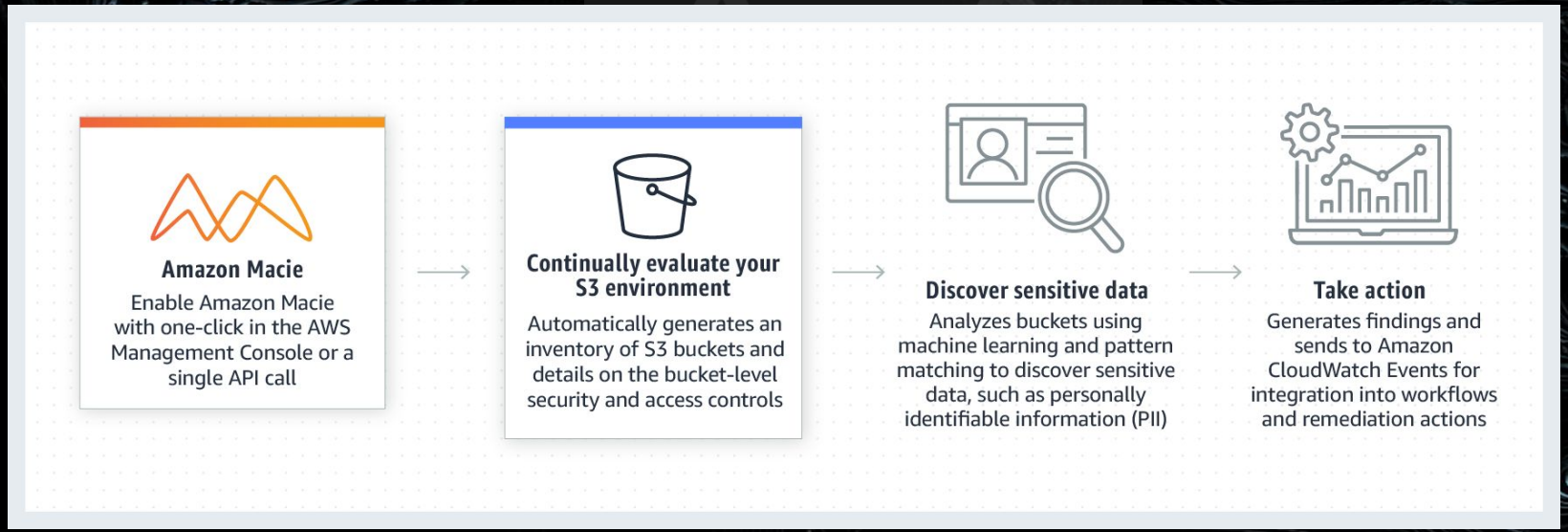
(Source: PII Data - AWS Macie Developer Guide)

$ Data Classification: AWS Well-Architected Framework

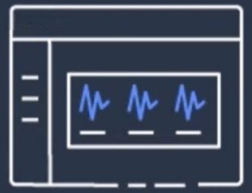
- Identify the data within your workload. (Associated business, compliance, etc.)
- Define data protection controls. (Sensitivity --> Tagging --> Control)
- Define data lifecycle management. (Sensitivity + Org's ISMS Controls --> Usable Approach)
- **Automate Identification and Classification**

$ Amazon Macie: Introduction

- 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. [Integrates with AWS S3]



$ Amazon Macie: Features



Gain visibility and evaluate

- Bucket inventory
- Bucket policies

Discover sensitive data

- Inspection jobs
- Flexible scope

Centrally manage at scale

- AWS Organizations
- Managed & custom data detections

Automate and take actions

- Detailed findings
- Management APIs

(Source: AWS Online Tech Talks)

$ Amazon Macie: Finding Types

Categories of Macie Findings

Policy Findings

- **Policy:IAMUser/S3BlockPublicAccessDisabled**
- **Policy:IAMUser/S3BucketEncryptionDisabled**
- **Policy:IAMUser/S3BucketPublic**
- **Policy:IAMUser/S3BucketReplicatedExternally**
- **Policy:IAMUser/S3BucketSharedExternally**

Sensitive Data Findings

- **SensitiveData:S3Object/Credentials**
- **SensitiveData:S3Object/CustomIdentifier**
- **SensitiveData:S3Object/Financial**
- **SensitiveData:S3Object/Multiple**
- **SensitiveData:S3Object/Personal**

$ Amazon Macie: Data Identifier Types

Sensitive Data Identifier Types

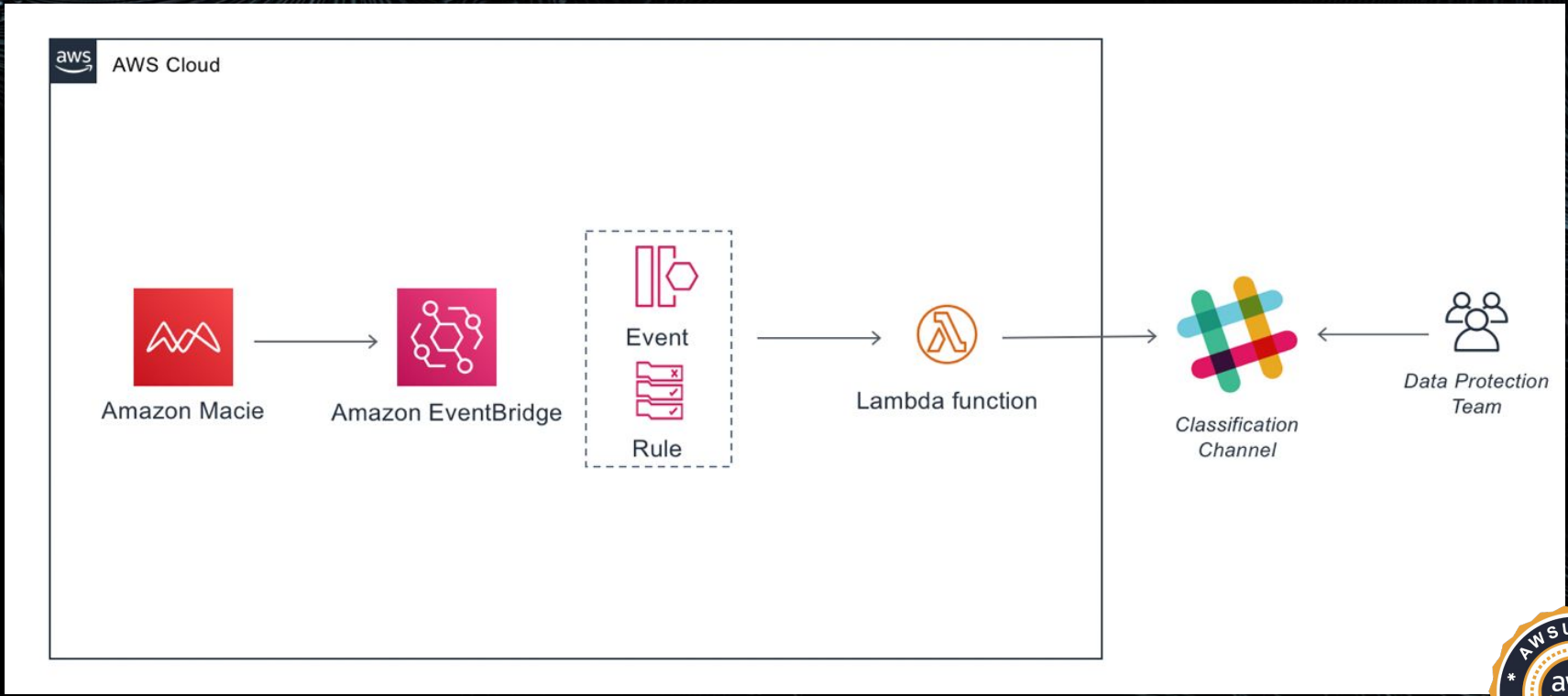
Fully Managed

- 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

Custom

- Regular expression that defines the pattern to match
- Keywords that define specific text to match
- Ignore words that define specific text to exclude

$ Amazon Macie: Demo - Sensitive Data Discovery



(Source: Sensitive Data Discovery with AWS Macie)

Thank You...

For queries feel free to connect with AWS Delhi User Group at:

1. **AWS Delhi User Group (MeetUp) - Learn Together, Grow Together**
2. **AWS User Group Delhi NCR (Follow us on LinkedIn Page)**