International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 13s, 2025
https://www.theaspd.com/ijes.php

Overview Of Data Classification And Applications In Data Security

Ha–Nam Nguyen[1], Le Cuong*[2]
[1, 2] Electric Power University, Hoang Quoc Viet Str., Bac Tu Liem District, Hanoi, Vietnam
namnhvn@epu.edu.vn[1], cuongle@epu.edu.vn[2]

Abstract
This paper provides a comprehensive overview of data classification frameworks, highlighting their significance in both governmental and commercial contexts. It explores the principles of data classification in the context of confidentiality, integrity, and availability (CIA), and outlines a systematic methodology for implementing and managing classification schemes. Key privacy regulations such as the GDPR, HIPAA, and CCPA are analyzed to underscore the importance of aligning classification practices with legal and compliance requirements. The study presents various models, including the U.S. National Security Classification Scheme and impact-based categorizations such as those defined by NIST. Additionally, the paper discusses emerging considerations in cloud computing environments and evaluates how cloud service providers support secure data classification and protection. Finally, it offers best practices for developing classification schemes that are context-aware, scalable, and compliant with evolving operational and regulatory needs.

Keywords: *Data classification, Cloud security, Information governance, Data privacy, cloud computing, data labeling.*

1. DATA CLASSIFICATION OVERVIEW

Data classification constitutes a fundamental component of cybersecurity risk management. It encompasses the systematic identification of data types that are processed and stored within an organization's information systems. Moreover, it entails the assessment of data sensitivity and the potential consequences associated with data compromise, loss, or misuse.To support effective risk mitigation, organizations are advised to adopt a context-driven approach to data classification. This involves analyzing the operational use cases of the data and developing a classification schema that reflects the potential impact on organizational functions—particularly in scenarios where data confidentiality, integrity, or availability is essential.Within the scope of this document, the term “classification” is employed in a comprehensive sense, encompassing the development of taxonomies, classification schemes, and categorization frameworks that align with the principles of confidentiality, integrity, and availability (CIA triad).

2. DATA CLASSIFICATION VALUE

Data classification has long served as a foundational practice for enabling organizations to apply appropriate levels of protection to sensitive or mission-critical information assets. Whether data resides in on-premises environments or is processed and stored in cloud infrastructures, classification is frequently regarded as a preliminary and essential step in determining the requisite controls to ensure the confidentiality, integrity, and availability of data, based on its associated risk to the organization.

For instance, information designated as confidential necessitates a higher degree of protection than publicly accessible data. Through data classification, organizations are empowered to assess the sensitivity and potential business impact of different data types, thereby facilitating a more nuanced evaluation of risk and enabling the application of proportional security measures.Standards bodies such as the International Organization for Standardization (ISO) and the National Institute of Standards and Technology (NIST) (ISO/IEC 27001, 2013; NIST, 2008) advocate for structured data classification frameworks. These frameworks assist in the effective governance and protection of information by aligning security controls with the relative sensitivity and criticality of the data. According to these standards, each classification tier should correspond to a baseline set of security controls designed to

mitigate vulnerabilities, counter threats, and manage risks in accordance with the level of protection warranted.However, it is important to acknowledge the potential drawbacks of over-classifying data. A common misstep among organizations is the indiscriminate assignment of high-sensitivity labels to heterogeneous datasets, which may result in unnecessary financial burdens due to the implementation of excessive security controls. Moreover, such over-classification can hinder operational efficiency, obscure the prioritization of truly critical assets, and impose restrictive compliance obligations that impede the optimal use of organizational data.

3. DATA CLASSIFICATION PROCESS

Organizations frequently require concrete and actionable guidance when formulating data classification policies. The following methodological framework, grounded in international standards and best practices, outlines a series of structured steps that not only support the initial development of classification schemes but also serve as evaluative criteria during the periodic reassessment of data classification accuracy and adequacy.

3.1. Establishing a data catalog

The initial phase involves constructing a comprehensive inventory of the various data types utilized and stored within the organization. This includes identifying the purposes for which the data is used, its sources, and determining whether it is subject to legal, regulatory, or internal compliance obligations. Once cataloged, data types should be systematically assigned to predefined classification tiers adopted by the organization. Tools such as AWS Glue Data Catalog can facilitate this process by enabling metadata storage, annotation, and sharing within cloud environments, while simultaneously offering audit capabilities, schema change tracking, and fine-grained access control.

3.2. Assessing Business-critical functions and conducting an impact assessment

A crucial determinant of appropriate data protection measures lies in understanding the business value and operational criticality of the data. Organizations should evaluate core business processes and map data dependencies accordingly. Following this assessment, an impact analysis should be conducted for each data category to estimate potential risks and consequences in the event of data compromise, loss, or misuse.

3.3. Labeling information

This step involves verifying and validating that data assets have been appropriately labeled in accordance with the organization's classification schema. A quality assurance process should be applied to ensure consistency and accuracy. Additionally, sub-labels may be introduced to address specific privacy or regulatory considerations within a broader classification tier. Technologies such as Amazon SageMaker and AWS Glue (Amazon Web Services, 2023) can be leveraged to support automated data labeling and enhance metadata granularity.

3.4. Handling of classified assets

Once classification has been assigned, the data must be managed in accordance with established handling protocols. These protocols prescribe security controls tailored to the sensitivity of each classification level, including requirements for encryption, access control, logging, and transmission. Handling procedures should be formalized in organizational documentation and periodically revised to remain aligned with evolving technologies and threat landscapes.

3.5. Continuous monitoring

To ensure the sustained integrity and effectiveness of the data classification program, organizations must implement mechanisms for continuous monitoring. These may include automated systems or manual reviews designed to observe security events, analyze access behaviors, apply patches and updates, and detect anomalies. Ongoing monitoring enables timely responses to emerging threats and facilitates the enforcement of data governance policies throughout the data lifecycle.

4. DATA CLASSIFICATION AND PRIVACY CONSIDERATIONS

Data classification has become increasingly critical in the context of emerging global privacy regulations, which afford individuals extensive rights over their personal information, including rights of access, rectification, deletion, and data portability. According to the United Nations Conference on Trade and

Development (UNCTAD) (UNCTAD, 2023), more than 70% of countries worldwide have enacted legislation pertaining to data protection and privacy, while an additional 10% are actively drafting such laws.For instance, the European Union's General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679, 2016) mandates that covered entities respond to certain data subject requests within one calendar month of receipt. Likewise, regulatory frameworks such as the California Consumer Privacy Act (CCPA) and the Health Insurance Portability and Accountability Act (HIPAA) (CCPA, 2018; HIPAA, 1996) grant individuals control over the collection, processing, and dissemination of their Personally Identifiable Information (PII) and Protected Health Information (PHI).To ensure compliance with these legal obligations, organizations must possess the capability to accurately verify the identity of the requester, locate and isolate the relevant personal data, guarantee that any disclosed data pertains exclusively to the requester, and appropriately deny requests that fall outside the scope of the applicable regulatory framework. In this context, the implementation of a robust data classification framework is indispensable. A well-structured data classification system-augmented by appropriate tagging and labeling mechanisms-facilitates the identification, protection, and retrieval of personal data. The use of sub-labels within broader classification tiers can further enhance the granularity of tagging, thereby enabling precise discovery of regulated data subsets. This capability not only supports timely and accurate fulfillment of data subject rights but also strengthens traceability and access control measures, ultimately promoting accountability in the management of sensitive information.

5. DATA CLASSIFICATION MODELS AND SCHEMES

Data classification models and schemes are generally categorized into two principal types: governmental classification schemes and commercial classification schemes. Governmental schemes are typically standardized and grounded in legal mandates, public policy frameworks, and executive directives. These models establish uniform criteria for data protection, often linked to national security or public interest considerations.In contrast, commercial classification schemes are inherently more flexible and organization-specific. They are developed based on the particular operational needs, risk tolerance, and sensitivity levels of the data handled within a given enterprise. Moreover, these schemes are often designed to align with industry-specific compliance obligations and regulatory requirements, such as those imposed by data protection laws or sectoral standards. Consequently, while governmental schemes emphasize uniformity and legal compliance, commercial schemes prioritize adaptability and contextual relevance in data protection strategies.

U.S. NATIONAL SECURITY CLASSIFICATION SCHEME

The United States government employs a three-tiered classification system for national security information, as established in Executive Order 13526 (Executive Order 13526, 2009). This framework is primarily concerned with the confidentiality dimension of information security, wherein classification levels are determined based on the projected severity of harm to national security resulting from unauthorized disclosure.

- **Confidential**: Information for which unauthorized disclosure could reasonably be expected to cause damage to national security.
- **Secret**: Information for which unauthorized disclosure could reasonably be expected to cause serious damage to national security.
- **Top Secret**: Information for which unauthorized disclosure could reasonably be expected to cause exceptionally grave damage to national security.

In addition to these primary classification levels, supplementary labels - commonly referred to as *caveats* - may be applied. These labels convey information regarding the origin of the data or impose additional handling constraints. Furthermore, the designation "Unclassified" is used to denote information not falling under the three primary classification levels. However, even unclassified data may be subject to additional controls. For example, categories such as *For Official Use Only (FOUO)* and *Controlled Unclassified Information (CUI)* are used to restrict disclosure to authorized personnel only, thereby acknowledging the sensitivity of such information despite its unclassified status.

U.S. INFORMATION CATEGORIZATION SCHEME (NIST FRAMEWORK)

Recognizing that national security classifications primarily address confidentiality risks, the National Institute of Standards and Technology (NIST) introduced an alternative categorization framework (FIPS PUB 199, 2004) to address a broader range of information security objectives. This framework, defined in Federal Information Processing Standards (FIPS) Publication 199, categorizes information and information systems based on the potential impact to the confidentiality, integrity, and availability of data in relation to an organization's mission and operations.

The three-tiered impact categorization includes:

- **Low**: The loss of confidentiality, integrity, or availability could be expected to have a limited adverse effect on organizational operations, assets, or individuals.
- **Moderate**: The loss could be expected to have a serious adverse effect on organizational operations, assets, or individuals.
- **High**: The loss could be expected to have a severe or catastrophic adverse effect on organizational operations, assets, or individuals.

This impact-based categorization model enables public sector organizations to implement risk-based security controls tailored to the criticality of their data and information systems across multiple security dimensions, rather than focusing exclusively on confidentiality. It is worth mentioning that in 2017, the city of Washington, D.C. enacted a new data classification policy aimed at enhancing transparency in government operations while simultaneously safeguarding sensitive information. The adopted framework comprises a five-tier classification model that, despite its granularity, maintains conceptual alignment with the widely used three-tier schemes commonly employed in cloud accreditation and information security regimes.

Level 0 – Open Data: This category includes datasets that are fully accessible to the public and are proactively published on open government platforms. The data in this tier is intended to promote transparency and civic engagement.

Level 1 – Public Data, Not Proactively Released: This refers to information that is not legally protected from public disclosure, nor subject to withholding under statutes, regulations, or contractual obligations. However, its publication may pose risks to privacy, safety, or security of individuals named in the data. Consequently, while not legally restricted, its dissemination is subject to cautious evaluation.

Level 2 – For District Government Use: This tier encompasses data that, while not highly sensitive, is intended for internal use within governmental entities. Such data is not restricted by legal or regulatory frameworks and primarily supports routine administrative functions and service delivery.

Level 3 – Confidential: This level includes data that is protected from disclosure by law, regulation, or contractual agreement. The information is considered sensitive due to its privacy-related content, such as Personally Identifiable Information (PII), Protected Health Information (PHI), Payment Card Industry Data Security Standard (PCI Security Standards Council, 2022) (PCI DSS) data, and Federal Tax Information (FTI). Unauthorized disclosure of such data could have serious legal and ethical implications.

Level 4 – Restricted Confidential: This highest classification level pertains to data whose unauthorized disclosure could result in significant harm, including threats to life or physical safety, or could substantially impair the operational capacity of governmental agencies to fulfill their statutory mandates. This tier demands the most stringent handling and access control protocols.

This five-level framework enables Washington, D.C. to balance openness and accountability with the critical need for data protection, and can serve as a referential model for jurisdictions seeking to implement nuanced data governance strategies.

6. INDUSTRY-SPECIFIC APPROACHES

This section presents industry-specific examples of data classification practices, which may incorporate sectoral requirements and regulatory mandates. As previously discussed, various data domains - such as governmental, financial, and healthcare information - often necessitate the application of specialized classification tiers and supplementary labels to ensure proper handling in accordance with legal, operational, and ethical standards.

Irrespective of whether the data is managed by public institutions or commercial entities, organizations are responsible for exercising due diligence to ensure compliance with applicable local, national, and international regulations. Data governance frameworks must be designed to reflect both the intrinsic sensitivity of the data and the contextual risks associated with its use, storage, and transmission.
The following table (not shown here) illustrates representative classification models currently implemented across different sectors. It outlines the characteristic features of each classification tier, provides examples of data types that fall within each category, and identifies corresponding workload or application types typically associated with each level of sensitivity. These practical examples serve as reference points for organizations developing or refining their own data classification strategies.

Table 1. Data classification – public sector

Data classification	Examples of workloads
Tier 3. Government confidential and above sensitive data	- National security and defense information - Government intelligence information - Law enforcement information - Government program monitoring or oversight investigations information
Tier 2. Restricted personally, identifying information about individuals	- Personally, identifying information about individuals - Human Resources Management - Personal profile information - Aggregated financial or market data
Tier 1. Public data	- Marketing or promotional information - Information related to other general government administrative or program activities - Intra-agency workplace policy development and management

7. COMMERCIAL DATA CLASSIFICATION SCHEME

Unlike governmental classification schemes - which are typically standardized and grounded in statutory mandates, regulatory frameworks, and executive directives - data classification models employed within commercial and non-governmental organizations are inherently more flexible and context-dependent. These schemes are typically tailored to the unique operational requirements and risk landscapes of individual organizations, particularly with respect to the sensitivity and business value of the data involved.Commercial classification schemes can vary significantly in complexity. Some organizations may adopt a basic two-tier model that distinguishes only between public and confidential data, while others may implement more granular, multi-level frameworks that incorporate nuanced distinctions among various data types and sensitivity levels.There exists no universally applicable methodology for designing a commercial data classification scheme. Instead, organizations are advised to evaluate their specific needs for protecting proprietary, operational, and user-generated data - each of which may vary in terms of confidentiality, integrity, and regulatory exposure. In developing an effective classification strategy, organizations should also account for compliance obligations under relevant legal and industry standards, as well as opportunities to align with prevailing cloud security best practices.Ultimately, a well-structured classification scheme should facilitate the systematic categorization of organizational data based on both criticality and sensitivity. Such categorization supports the implementation of appropriate protection measures, including access controls, encryption, and retention policies, thereby contributing to a comprehensive and risk-informed information governance strategy.

Table 2 - Five-tiered commercial data classification approach according to the book CISSP Security Management and Practices

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 kept 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

We recommends using the minimal number of tiers that make sense for the organization as follow:
Table 3 — Three-tiered data classification approach

Data classification	System security categoriz action	Cloud deployment model options
Unclassified	Low to High	Accredited public cloud
Official	Moderate to High	Accredited public cloud
Secret and above	Moderate to High	Accredited private/hybrid/community cloud/public cloud

Data Residency Considerations: As part of a comprehensive data governance strategy, organizations are encouraged to critically evaluate their data classification frameworks with respect to data residency requirements. This involves determining which categories of data must remain within specific geographic boundaries - such as national or regional jurisdictions - and articulating the underlying legal, regulatory, or operational justifications for such restrictions.In the absence of explicit statutory or contractual obligations mandating data localization, organizations may discover that even sensitive or mission-critical data can be lawfully stored or replicated in other jurisdictions. Leveraging cross-regional data storage capabilities can yield several strategic benefits, including enhanced resilience through disaster recovery failover mechanisms and access to advanced technologies or services not available locally.Accordingly, data residency planning should not only reflect compliance obligations, but also incorporate a risk-based assessment of operational continuity, data sovereignty, and jurisdictional exposure. Such analysis is essential for balancing regulatory adherence with performance optimization and organizational agility in distributed computing environments.Organizations are optimally positioned to design and implement data classification schemes that are tailored to their unique operational contexts and risk management priorities. Rather than adhering rigidly to complex, multi-tiered classification frameworks, entities may conduct structured risk impact assessments to determine whether simplified models - such as a three-tier scheme - would better align with their data governance and operational requirements.The selection of an appropriate cloud deployment model should be informed by the organization's specific use cases, the nature and sensitivity of the data involved, and the outcomes of a formal risk assessment. Based on the assigned data classification level, organizations must implement corresponding security controls - such as encryption, access restrictions, and logging - within the cloud environment to ensure data confidentiality, integrity, and availability.In assessing risk and defining control requirements, it is imperative to recognize the fundamental differences between cloud-based and on-premises infrastructures. This includes understanding alternative control mechanisms, the nuances of control implementation, and the implications of the cloud provider's shared responsibility model. Such considerations are essential to

ensuring that appropriate safeguards are applied consistently across all service delivery models.Upon comprehensive evaluation of commercial cloud services and their associated security advantages - including improved service availability, operational resilience, enhanced visibility and automation, and continuous security auditing - many organizations may conclude that a substantial portion of their workloads can be securely migrated to the cloud. This is particularly viable when guided by a well-defined data classification framework, as demonstrated by initiatives led by governments such as those of the United States and the United Kingdom.Globally, public sector organizations are increasingly capitalizing on the inherent security features of commercial cloud platforms. By coupling these capabilities with robust data classification policies and the implementation of context-appropriate security controls, these entities are advancing toward enhanced compliance, security posture, and service efficiency in cloud-native environments.

8. EMERGING CONSIDERATIONS IN DATA CLASSIFICATION

Irrespective of whether an organization is newly transitioning to cloud environments or has an established cloud infrastructure, the formulation and continuous refinement of data classification policies remains a critical component of a robust data governance strategy. Similar to the periodic reassessment of security protocols in response to evolving threat landscapes, data classification frameworks must also be revisited to reflect emerging technological, organizational, and operational dynamics.Recent discourse within industry consortiums and professional forums has brought attention to several contemporary challenges that organizations should address when reviewing or updating their data classification policies:

Data is scattered everywhere: The pervasive adoption of digital technologies and the increasing reliance on data-driven operations have resulted in the dispersion of vast volumes of data across diverse platforms, devices, and endpoints. The decentralized nature of modern data ecosystems presents significant challenges for organizations in maintaining visibility, control, and consistent protection across the entire data lifecycle.

Intra- and Inter-Organizational Dependencies: Collaborative workflows, both within and between organizations -particularly in sectors such as healthcare, government, and critical infrastructure - necessitate the frequent exchange and shared management of sensitive information. These dependencies underscore the need for harmonized classification policies and interoperable security protocols.

End-User Knowledge: Classification models that depend on manual input from end users (e.g., for purposes of machine learning or document tagging) are susceptible to inconsistencies and inaccuracies. End users may lack the necessary training or risk awareness to accurately assess the sensitivity of the data they handle, leading to under- or over-classification.

Data classifiers and tagging: The absence of standardized taxonomies or definitions for classification labels within certain industries contributes to ambiguity in implementation. Furthermore, inconsistent tagging practices may impede downstream processes such as data retrieval, access control, and compliance auditing.

Context: The true sensitivity or criticality of information is often context-dependent. Beyond the intrinsic content of the data, its value and risk profile are influenced by factors such as how the data is used, with whom it is shared, and the potential consequences of unauthorized disclosure. Context-aware classification mechanisms are therefore essential for accurate data protection.

Although these considerations are not entirely novel, they are increasingly salient in light of modern computing environments and evolving data practices. Organizations should integrate these emerging factors into the design and implementation of data classification schemes to ensure that policies remain effective, scalable, and aligned with both operational and regulatory requirements.

9. USING CLOUD SERVICES TO SUPPORT DATA CLASSIFICATION

Cloud computing can provide customers with the ability to secure their workloads. Organizations in highly regulated industries, the public sector, enterprises, small and medium-sized businesses, or startups can work to meet their data classification policies and requirements in the cloud. Cloud service providers (CSPs), such as AWS, provide services based on standardized utilities that customers provide themselves.

The cloud does not have visibility into the type of data that customers run in the cloud, which means that the cloud does not differentiate, for example, personal data from other customer data when providing cloud services. It is the customer's responsibility to classify their data and implement appropriate controls in their cloud environment (e.g. encryption). However, the security controls that CSPs implement in their infrastructure and services can be used by customers to help meet their most sensitive data requirements.
Cloud Conputing Security and Compliance: Cloud computing services provide the same high level of security to all customers, regardless of the type of content stored. These services are then aligned to international "***gold***" standards for security and compliance, meaning customers benefit from a higher level of protection for customer data processed and stored in the cloud.The most concerning risk events and threat vectors are largely addressed through basic cyber hygiene principles (such as patching and system configuration), which CSPs can demonstrate through globally recognized security certifications and assurance programs such as ISO 27001 (ISO/IEC 27001, 2013), the Payment Card Industry Data Security Standard (PCI Security Standards Council, 2022) (PCI DSS), and Service Organization Controls (SOC). ISO 27001/27002 is a widely adopted global security standard that sets out requirements and best practices for a systematic approach to managing corporate and customer information based on periodic risk assessments that are consistent with ever-changing threat scenarios.The Payment Card Industry Data Security Standard (PCI Security Standards Council, 2022) (also known as PCI DSS) is a proprietary information security standard managed by the PCI Security Standards Council, which was established by American Express, Discover Financial Services, JCB International, MasterCard Worldwide, and Visa Inc. PCI DSS applies to organizations that store, process, or transmit card data.
Service Organization Control Reports (AICPA, 2022) (SOC 1, 2, 3) are intended to meet a variety of financial audit requirements for U.S. and international auditing agencies. Audits for these reports are conducted in accordance with International Standard for Assurance Engagements 3402 (ISAE 3402) and the American Institute of Certified Public Accountants (AICPA): AT 801 (formerly SSAE 16).Security and compliance reports such as SOC 1, PCI, FedRAMP are available to customers through a self-service portal for on-demand access to security and compliance reports for each cloud service type. You can use these documents to validate the implementation and effectiveness of security controls on your cloud services. These documents can also be used as a guide to assess and validate the effectiveness of your company's internal controls. Cloud computing customers are responsible for developing or obtaining documentation that demonstrates the security and compliance of their workloads on the cloud service. For more information, refer to the Shared Responsibility Model (Amazon Web Services, 2023). When evaluating CSPs, organizations should leverage these existing CSP certifications so they can accurately determine whether the CSP (and the services within the CSP's offerings) can support their data classification requirements. Cloud computing services encourages organizations to implement a policy that identifies which existing national, international, or industry-specific cloud certifications and certifications are acceptable for each level of their data classification program to streamline recognition and expedite the migration of workloads to the cloud. The security pillar provides guidance to help you apply best practices and current recommendations in designing, provisioning, and maintaining secure cloud workloads.

Two of the design principles that focus on data protection include:

- **Protect data in transit and at rest** - Classify your data into sensitivity levels and use mechanisms such as encryption, tokenization, and access controls where appropriate.
- **Keep people away from data** - Use mechanisms and tools to reduce or eliminate the need for direct access or manual handling of data. This reduces the risk of mishandling or modification and human error when handling sensitive data.

In terms of data classification, this work provides the following additional recommendations:

- **Identify the data in your workloads** - Understand the type and classification of data your workloads are processing, the business processes involved, the data owners, applicable legal and compliance requirements, where the data is stored, and the controls that need to be enforced.
- **Identify data protection controls** - Using resource tags, separate cloud service accounts by sensitivity (and potentially also by alert/protection zone/community of interest), Identity and Access

Management (IAM) policies, organizational Service Control Policies (SCPs), cloud service Key Management Services, and related services, organizations can define and implement policies to classify and protect data.

- **Identify data lifecycle management** - Have a defined lifecycle strategy based on sensitivity and legal and organizational requirements. Consider how long your organization must retain data, how it destroys data, how it manages access to data, how it transforms data, and how it shares data.
- **Automated identification and classification** - Automating the identification and classification of data, as opposed to directing access from an individual or group, reduces the risk of human error/exposure and helps implement accurate controls.

Cloud computing services and its features: Cloud computing services provide a number of services and features that can facilitate the implementation of an organization's data classification program. For example, Amazon Macie can help customers catalog and classify sensitive and business - critical data stored in AWS. Amazon Macie uses ML to automate the process of discovering, classifying, labeling, and applying protection rules to data. This helps customers better understand where sensitive information is stored and how it is accessed, including user authentication and access patterns. Another important feature that supports data classification and protection is cloud resource tagging. By assigning metadata to resources in your cloud services in the form of tags (each tag is a label consisting of a user-defined key and value), you can manage, identify, sort, search, and filter resources. Security tags can contain security information, identifying the specific data security levels that a resource supports or complies with, such as identifiers for workloads that must comply with specific compliance requirements.

10. CONCLUSION

This work has examined data classification frameworks across both governmental and commercial sectors, emphasizing key distinctions in categorization practices and their alignment with regulatory and operational objectives. It has also addressed critical privacy considerations, outlined essential factors for implementing effective classification schemes, and provided guidance for reviewing, developing, and refining existing data classification policies.In addition, the paper has presented a set of best practices and strategic recommendations for leveraging cloud computing services to support classification initiatives. These include methods to align classification processes with data sensitivity requirements and ensure that appropriate security controls are deployed to protect organizational data in distributed cloud environments.

FUNDING

This research is funded by Electric Power University under research 2025.

REFERENCES

1. Executive Order 13526 (Executive Order 13526, 2009). (2009). Classified National Security Information. Federal Register, 75(2).
2. FIPS PUB 199. (2004). Standards for Security Categorization of Federal Information and Information Systems. National Institute of Standards and Technology.
3. ISO/IEC 27001. (2013). Information technology - Security techniques - Information security management systems - Requirements.
4. ISO/IEC 27002. (2022). Information security, cybersecurity and privacy protection - Information security controls.
5. NIST. (2008). Guide for Mapping Types of Information and Information Systems to Security Categories (SP 800-60 Rev.1).
6. Regulation (EU) 2016/679. (2016). General Data Protection Regulation (GDPR). Official Journal of the European Union.
7. California Consumer Privacy Act (CCPA). (2018). California Civil Code, Section 1798.100.
8. Health Insurance Portability and Accountability Act (HIPAA). (1996). U.S. Department of Health & Human Services.
9. UNCTAD. (2023). Data Protection and Privacy Legislation Worldwide. United Nations Conference on Trade and Development.
10. Amazon Web Services. (2023). Amazon Macie User Guide. Retrieved from https://docs.aws.amazon.com/macie/
11. Amazon Web Services. (2023). Shared Responsibility Model (Amazon Web Services, 2023). Retrieved from https://aws.amazon.com/compliance/shared-responsibility-model/
12. PCI Security Standards Council. (2022). PCI DSS v4.0 Documentation.
13. AICPA. (2022). SOC Reports - System and Organization Controls. Retrieved from https://www.aicpa.org/