
Ordis

AI DATA PROTECTION AND CLASSIFICATION POLICY

Document Title	**AI Data Protection and Classification Policy**
Document ID	
Version	**1.0**

Department	Information Security / GDPR
Unit	Information Systems, AI Systems, Compliance, Legal, Risk Management and Governance
Approved By	John Dabill
Applies to	The organization's employees, contractors/consultants, temporary employees and third parties.
Effective Date	*13th August 2026*
Review Date	*13th August 2027*

CONTENTS

1. PURPOSE

With the evolving technology of artificial intelligence (AI) the need for efficient governance, ongoing assessment, and user awareness arises. A vital component of understanding how to safeguard AI-related data and systems is the classification of AI systems. A classification framework for AI systems supports the distinguishing characteristics of AI systems and provides a structured way to breakdown this technology. AI system classification is not as specific as other data classification frameworks; it is populated through the input AI systems stakeholders.

2. SCOPE

This policy applies to all employees, contractors, consultants, temporary staff, third-party users, and any other stakeholders who have access to the organization's information systems, data and applications. It applies to all forms of access, including physical access, logical access, and remote access, and to all types of devices, including desktops, laptops, mobile devices, and servers.

3. ROLES AND RESPONSIBILITIES

3.1. Data Owners/ Controllers

- Responsible for the data throughout its lifecycle.
- Define the classification levels and criteria for data based on its sensitivity, value, and regulatory requirements.
- Assign the appropriate classification level to data and ensure it is properly labeled or tagged.
- Authorize access to data based on the classification level and the need-to-know principle.
- Regularly review the classification of data to ensure it is accurate and up to date.
- Establish data governance policies and procedures to guide the classification process and ensure compliance with relevant regulations.

3.2. Data Custodians/ Processors

- Responsible for the storage, protection, and maintenance of data.
- Implement appropriate security controls, such as encryption, access controls, and backups, to protect classified data.
- Ensure that classified data is stored securely, adhering to retention and disposal policies.
- Participate in incident response activities, including detecting and responding to data breaches or unauthorized access to classified data.
- Monitor compliance with data classification policies and report any non-compliance issues to data owners.

- Ensure the integrity and quality of classified data by implementing data validation and verification processes.

3.3. End Users

- Users refer to individuals or groups who access and utilize classified data.
- Must understand and comply with data classification policies, including handling and protection requirements for each classification level.
- Responsible for accessing and using classified data only as authorized based on their roles and responsibilities.
- Ensure that classified data is shared securely and in accordance with applicable data sharing agreements or protocols.
- Report any incidents, breaches, or concerns related to classified data to the appropriate authorities or data custodians.
- Participate in training programs to enhance their understanding of data classification principles and best practices.

4. ARTIFICIAL INTELLIGENCE SYSTEMS CLASSIFICATION

AI systems can be classified based on different criteria, including their capabilities, functionality, and level of autonomy. It is important to note that AI system classifications can overlap, and many AI systems may incorporate multiple techniques and approaches. The classification of AI systems may evolve as technology advances and new capabilities emerge.

4.1. Narrow AI

Designed to perform specific tasks or functions within a limited domain. Focused on solving specific problems and excel in narrow areas such as speech recognition, image classification, or recommendation systems.

4.2. General AI

AI systems that possess human-like intelligence and can understand, learn, and apply knowledge across various domains. General AI can perform tasks that require human-level cognitive abilities and adapt to new situations.

4.3. Machine Learning (ML) Systems

Systems that use algorithms and statistical models to learn from data and make predictions or decisions. ML systems analyze large datasets, identify patterns, and improve their performance over time through iterative learning processes. Machine learning can be further classified into supervised learning, unsupervised learning, and reinforcement learning based on the type of training data and learning methods used.

4.4. Deep Learning Systems

Systems that are a subset of machine learning systems that utilize artificial neural networks with multiple layers to learn and extract complex patterns from data, such as computer vision, natural language processing, and speech recognition.

4.5. Autonomous AI Systems

Systems that can make decisions and take actions without direct human intervention, with a high degree of autonomy and operating independently in real-world environments.

4.6. Expert Systems

Designed to mimic the expertise and decision-making capabilities of human experts in specific domains. These systems use a knowledge base and inference engines to analyze data and provide recommendations or solutions.

AI System Classification	Impact			Autonomy	
Narrow AI	**Low**	**Medium**	**High**	**Low**	**Medium**
General AI	**High**			**High**	
Machine Learning (ML)	**High**			**High**	
Deep Learning	**High**			**Medium**	**High**
Autonomous AI	**High**			**High**	
Expert Systems	**High**			**Medium**	**High**

Table 1: AI System Classification

5. ARTIFICIAL INTELLIGENCE APPLICATIONS

AI applications can be classified based on several factors, including their functionality, problem-solving approach, and domain of application. Below the most common classifications of AI applications are defined:

5.1. Task-Specific AI

Applications that have been designed to perform a specific task, focused on solving a particular problem or addressing a specific need (e.g., voice assistants, image recognition systems, and chatbots).

5.2. General-Purpose AI or Artificial General Intelligence (AGI)

AGI systems are capable of understanding, learning, and performing tasks that a human can do. These applications aim to exhibit human-level intelligence across a wide range of tasks and domains.

5.3. Vertical AI

Applications tailored to specific industries or domains that address the unique challenges and requirements of a particular sector (e.g., AI applications in healthcare, finance, manufacturing, and transportation).

5.4. Horizontal AI

Applications that have a broad applicability and can be used across different industries and domains, providing AI capabilities that can be applied to various tasks and functions (e.g., natural language processing, machine vision, predictive analytics).

5.5. Reactive AI

Applications that do not possess the ability to learn from experience and can only react to specific inputs or stimuli. Reactive AI systems do not possess the capability to retain past information or adjust their behavior based on previous interactions.

5.6. Limited Memory AI

Applications that can learn from historical data and make decisions based on past experiences. However, they have a limited form of memory that allows them to improve their performance over time (e.g., recommendation systems, some machine learning algorithms).

5.7. Self-Aware AI

Applications that possess self-awareness and consciousness, understanding their own existence and the surrounding environment. Self-aware AI systems have an advanced level of cognition and the ability to introspect.

Type of AI Application	Impact			Autonomy	
Task-Specific AI	Low	Medium	High	Low	Medium
General-Purpose AI	High			Medium	High
Vertical AI	High			Medium	High
Horizontal AI	High			Medium	High
Reactive AI	Low			Low	
Limited Memory AI	Medium			Medium	
Self-Aware AI	High			High	

Table 2: Types of AI Applications

6. CLASSIFICATION LEVELS FOR AI SYSTEMS AND APPLICATIONS

Evaluating the following criteria and factors allows the organization to determine the appropriate classification level for each AI data type. This ensures that adequate protection measures are applied based on the level of sensitivity and potential risk associated with the data. The classification levels used for AI systems, data and applications are as follows:

6.1. Public

Data classified as Public is intended for unrestricted and public distribution. This category typically includes information that poses no risk or harm if disclosed to the public. Examples may include publicly available research papers, non-sensitive marketing materials, or general information about the organization.

6.2. Internal

Data classified as Internal is intended for internal use within the organization. This category includes information that is not intended for public disclosure but may not contain sensitive or confidential details. Examples may include internal memos, non-sensitive employee information, or routine operational data.

6.3. Confidential

Data classified as Confidential is sensitive and requires protection from unauthorized access or disclosure. This category includes information that, if accessed by unauthorized individuals, could result in harm to individuals, the organization, or stakeholders. Examples may include customer information, financial data, intellectual property, or proprietary algorithms.

6.4. Highly Confidential

Data classified as Highly Confidential is the highest level of sensitivity and requires the utmost protection. This category includes information that, if accessed or disclosed improperly, could cause significant harm to individuals, the organization, or stakeholders. Examples may include personal health information, financial account credentials, trade secrets, or classified research data.

Classification Level	Impact	Autonomy	Purpose or Risk
Public	**Low**	**Low**	Data poses no risk if disclosed to the public.
Internal	**Low** / **Medium**	**Medium**	Intended for internal use within the organization.
Confidential	**Medium** / **High**	**High**	Sensitive information requiring protection from unauthorized access or disclosure.
Highly Confidential	**High**	**High**	Highest level of sensitivity requiring utmost protection. Potential for significant harm if accessed or disclosed improperly.

Table 3:AI Data Classification Levels

The determination of the appropriate classification level for each AI data type involves considering several criteria and factors, including:

- ★ **Data Sensitivity** - The level of sensitivity of the data and the potential impact if accessed or disclosed improperly.
- ★ **Legal and Regulatory Requirements** - Compliance with applicable laws, regulations, and industry standards related to data protection and privacy.
- ★ **Data Ownership and Control** - The ownership of the data and the level of control required to protect it adequately.
- ★ **Data Value** - The value of the data to the organization and the potential consequences of its loss, theft, or unauthorized disclosure.
- ★ **Data Usage and Access** - The intended usage of the data and the level of access required by different individuals or groups within the organization.
- ★ **Risk Assessment** - Assessing the potential risks and threats associated with the data and the likelihood of those risks occurring.
- ★ **Data Retention Requirements** - Considering the required retention period for the data and the need for secure storage during that period.

7. HANDLING AND PROTECTION REQUIREMENTS FOR AI DATA

The organization has implemented data handling and protection procedures to specify the requirements based on the AI data classification level. This may include access controls, encryption, data storage guidelines, and disposal procedures.

Classification Level	Access Controls	Encryption	Data Storage Guidelines	Disposal Procedures
Public	Data should be accessible to the public without restrictions.	No encryption requirements.	Data can be stored in publicly accessible systems.	No special disposal procedures required.
Internal	Limited access to authorized personnel within the organization.	Encryption may be required for data transmission and storage.	Data should be stored in secure internal systems with access controls.	Disposal should follow proper data destruction methods.
Confidential	Strict access controls should be implemented, granting access only to authorized	Encryption should be applied to data both at rest and in transit.	Data should be stored in secure systems with strong access controls and monitoring.	Disposal should follow secure data destruction methods, such as secure wiping or

Classification Level	Access Controls	Encryption	Data Storage Guidelines	Disposal Procedures
	individuals on a need-to-know basis.			physical destruction.
Highly Confidential	Access should be limited to a small group of authorized individuals with a legitimate need.	Strong encryption should be applied to data both at rest and in transit.	Data should be stored in highly secure systems with advanced access controls, intrusion detection, and monitoring.	Disposal should follow stringent procedures, such as physical destruction under supervision and documented processes.

Table 4: AI Data Handling and Protection Procedures

8. DATA CLASSIFICATION PROCEDURES

The organization has established procedures to effectively classify their data, ensure appropriate handling and protection, and meet regulatory and compliance requirements. The following steps are involved in the classifying data procedures:

8.1. Identification of Data Owners

8.1.1. Identify and assign data owners for different data sets or data categories.

8.2. Assessment of Data Sensitivity

8.2.1. Conduct a data sensitivity assessment to determine the sensitivity level of the data.

8.2.2. Consider factors such as legal requirements, privacy concerns, potential impact on individuals or the organization, and contractual obligations.

8.2.3. Involve relevant stakeholders, such as legal, compliance, and information security teams, in the assessment process.

8.3. Classification Assignment

8.3.1. Based on the data sensitivity assessment, assign an appropriate classification level (e.g., Public, Internal, Confidential, Highly Confidential) to each data set or data category.

8.3.2. Consider the impact and autonomy factors specific to AI data types to determine the classification level (Low, Medium, or High).

8.4. Documentation

8.4.1. Document the data classification decisions and the assigned classification levels for

each data set or data category.

8.4.2. Maintain a central repository or database where the classification information is recorded.

8.4.3. Ensure the documentation is regularly updated to reflect any changes or updates in the classification of data.

8.5. Communication and Awareness

8.5.1. Communicate the data classification framework and the assigned classification levels to relevant stakeholders, including data owners, data custodians, and users.

8.5.2. Raise awareness about the importance of data classification, handling, and protection among employees.

8.5.3. Provide training and guidance on the proper handling and protection of data based on its assigned classification level.

8.6. Periodic Review and Reassessment

8.6.1. Conduct periodic reviews and reassessments of data classification to ensure its ongoing relevance and accuracy.

8.6.2. Consider changes in regulations, business requirements, and the evolving nature of data to update the classification as needed.

8.6.3. Involve data owners and relevant stakeholders in the review process to gather insights and ensure alignment with organizational needs.

9. AI DATA INCIDENT RESPONSE

The organization has implemented procedures to effectively respond to AI incidents and mitigate the impact. Additionally, the incident response procedures work towards strengthening their security posture and preventing future incidents. In the event of an AI data breach or unauthorized access to AI classified data, the following steps should be followed:

9.1. Detection and Initial Response

9.1.1. Detect and confirm the occurrence of a data breach or unauthorized access.

9.1.2. Immediately initiate the incident response process and notify the appropriate personnel or teams.

9.2. Containment and Isolation

9.2.1. Take immediate steps to contain and isolate the breach or unauthorized access.

9.2.2. Disable compromised accounts or systems, if necessary, to prevent further damage or unauthorized activities.

9.3. Reporting

9.3.1. Follow established reporting procedures and notify the designated individuals or teams responsible for managing data breaches.

9.3.2. Provide a detailed incident report, including the nature of the breach, the data affected, and the initial assessment of the impact.

9.4. Investigation

9.4.1. Conduct a thorough investigation to determine the cause, scope, and extent of the data breach or unauthorized access.

9.4.2. Preserve evidence, collect relevant logs and records, and document the findings throughout the investigation process.

9.5. Risk Assessment and Impact Analysis

9.5.1. Assess the potential risks and impacts resulting from the data breach or unauthorized access.

9.5.2. Determine the severity and likelihood of harm to affected individuals, systems, or the organization.

9.5.3. Identify any legal or regulatory obligations that must be fulfilled.

9.6. Notification and Communication

9.6.1. Comply with any legal or regulatory requirements for notifying affected individuals or authorities about the data breach.

9.6.2. Communicate transparently and effectively with affected parties, providing clear information about the breach, the potential risks, and any steps they should take to protect themselves.

9.7. Remediation and Mitigation

9.7.1. Take immediate steps to remediate the breach or unauthorized access and prevent further damage.

9.7.2. Implement appropriate measures to secure the affected systems, data, or vulnerabilities.

9.7.3. Update security controls, procedures, or policies as necessary to prevent similar incidents in the future.

9.8. Evaluation and Lessons Learned

9.8.1. Conduct a post-incident review to evaluate the organization's response to the data breach or unauthorized access.

9.8.2. Identify any weaknesses or gaps in security controls, incident response procedures,

or employee awareness.

9.8.3. Use the lessons learned to improve incident response capabilities and enhance data protection measures.

9.9. Documentation and Record-Keeping

9.9.1. Document all actions, findings, and remediation efforts throughout the incident response process.

9.9.2. Maintain and retain records of the incident, including investigation reports, communication logs, and any relevant evidence or documentation.

9.9.3. Adhere to data retention policies and regulatory requirements for storing incident-related records.

10. MONITORING OF POLICY COMPLIANCE

The organization has implemented monitoring and enforcement measures to ensure that compliance with the data classification policy is upheld and data is appropriately classified, protected, and managed throughout its lifecycle. Compliance with the data classification policy will be monitored and enforced through the following measures:

10.1. Regular Audits and Assessments

10.1.1. Periodic audits and assessments to evaluate the implementation of the data classification policy.

10.1.2. Use of internal or external auditors to assess the adherence to the policy's requirements, including the proper classification of data, access controls, encryption practices, and data handling procedures.

10.1.3. Identify any gaps or non-compliance issues and take corrective actions to address them.

10.2. Access Control Mechanisms

10.2.1. Implement robust access control mechanisms that align with the data classification levels.

10.2.2. Utilize role-based access controls (RBAC) to ensure that employees only have access to data that is appropriate for their roles and responsibilities.

10.2.3. Regularly review and update user permissions and access rights to align with changes in roles or job responsibilities.

10.3. Data Handling Monitoring

10.3.1. Implement monitoring mechanisms to track and log data access, usage, and modifications.

10.3.2. Use security technologies such as data loss prevention (DLP) systems or activity monitoring tools to detect and prevent unauthorized access or data breaches.

10.3.3. Monitor the movement of data within the organization, including transfers between systems or external parties, to ensure compliance with the data classification policy.

10.4. Training and Awareness Programs

10.4.1. Continuously educate employees about the data classification policy, its importance, and the consequences of non-compliance.

10.4.2. Provide regular training sessions and awareness programs to reinforce understanding and promote adherence to the policy.

10.4.3. Incorporate data classification compliance as part of employee onboarding processes to ensure new hires are aware of their responsibilities.

10.5. Continuous Improvement

10.5.1. Regularly review and update the data classification policy to reflect changes in technology, regulations, or organizational requirements.

10.5.2. Collect feedback from stakeholders and employees to identify areas for improvement and address any challenges or gaps in the policy implementation.

10.5.3. Foster a culture of continuous improvement and encourage suggestions for enhancing data classification practices.

11. POLICY COMPLIANCE

11.1. Compliance Measurement

The organisation will verify compliance with this policy through various methods, including but not limited to, business tool reports, internal and external audits, and feedback to the policy owner.

11.2. Exceptions

Any exceptions to the policy must be previously approved and formally documented by Senior Management.

11.3. Non-Compliance

Any member of the company's personnel found to have violated this policy may be subject to disciplinary actions, up to and including termination of employment.

12. DEFINITIONS

a. **AI Data** - Any data generated, processed, or used by artificial intelligence systems, including but not limited to machine learning models, deep learning algorithms, expert systems, and autonomous AI systems.

b. **Data Classification** - The process of categorizing data based on its sensitivity, value, and risk level to determine appropriate handling, storage, and protection measures.

c. **Data Owner** - The individual or entity responsible for the overall governance and accountability of specific data assets. The data owner is responsible for determining the classification level and ensuring appropriate security controls are in place.

d. **Data Custodian** - The individual or entity responsible for the day-to-day management and protection of data, including implementing and maintaining the security controls specified by the data owner.

e. **Data User** - Any individual or entity authorized to access and use classified data for legitimate business purposes. Data users must adhere to the data classification policy and follow the specified security controls.

f. **Data Sensitivity** - The level of sensitivity or confidentiality associated with the data, considering factors such as legal requirements, privacy concerns, intellectual property, and business impact.

g. **Classification Levels** - The predefined categories used to classify data based on its sensitivity and criticality. The classification levels may include *Public*, *Internal*, *Confidential*, and *Highly Confidential*, with each level having specific handling and protection requirements.

h. **Impact** - The potential consequences or risks associated with unauthorized access, disclosure, or modification of data. Impact levels may include Low, Medium, or High, reflecting the potential harm or damage that may result from a security incident.

i. **Autonomy** - The level of independence or decision-making capability of an AI system. Autonomy levels may include Low, Medium, or High, indicating the system's ability to operate and make decisions without human intervention.

j. **Security Controls** - Measures, policies, and procedures implemented to protect classified data, including access controls, encryption, data storage guidelines, disposal procedures, and monitoring mechanisms.

k. **Data Classification Policy** - The formal document that outlines the principles, procedures, and guidelines for classifying data, determining its handling and protection requirements, and assigning ownership and responsibility for data management and security.

l. **Non-Disclosure Agreement (NDA)** - A legal agreement between parties that outlines the terms and conditions for the protection and confidentiality of sensitive information, including classified data.

13. REVISIONS AND APPROVALS

13.1. Document History

Version	Date	Description	Comments
1.0	xxx	Document creation.	

13.2. Approvals by Level

Approver	Approver Title	Date Approved	Comments