Data Classification Policy

Guidelines for Data Classification

Purpose

The purpose of this Guideline is to establish a framework for classifying data based on its level of sensitivity, value and criticality to Thematic as required by Thematic's Information Security Policy. Classification of data will aid in determining baseline security controls for the protection of data.

Applies To

This Policy applies to all staff and third-party Agents of Thematic as well as any other Thematic affiliate who is authorized to access Thematic Data. In particular, this Guideline applies to those who are responsible for classifying and protecting Thematic Data, as defined by the Information Security Roles and Responsibilities

Definitions

Confidential Data is a generalized term that typically represents data classified as Restricted, according to the data classification scheme defined in this Guideline. This term is often used interchangeably with sensitive data.

A *Data Steward* is a senior-level employee of Thematic who oversees the lifecycle of one or more sets of Thematic Data.

Thematic Data is defined as all data owned or licensed by Thematic.

Non-public Information is defined as any information that is classified as Private or Restricted Information according to the data classification scheme defined in this Guideline.

Sensitive Data is a generalized term that typically represents data classified as Restricted, according to the data classification scheme defined in this Guideline. This term is often used interchangeably with confidential data.

Data Classification

Data classification, in the context of information security, is the classification of data based on its level of sensitivity and the impact to Thematic should that data be disclosed, altered or destroyed without authorization. The classification of data helps determine what baseline security controls are appropriate for safeguarding that data. All Thematic data should be classified into one of three sensitivity levels, or classifications:

A.	**Restricted Data**
	Data should be classified as Restricted when the unauthorized disclosure, alteration or destruction of that data could cause a significant level of risk to Thematic, its customers or its affiliates. Examples of Restricted data include data protected by state or federal privacy regulations and data protected by confidentiality agreements. The highest level of security controls should be applied to Restricted data.
B.	**Private Data**
	Data should be classified as Private when the unauthorized disclosure, alteration or destruction of that data could result in a moderate level of risk to Thematic, its customers or its affiliates. By default, all Thematic Data that is not explicitly classified as Restricted or Public data should be treated as Private data. A reasonable level of security controls should be applied to Private data.
C.	**Public Data**
	Data should be classified as Public when the unauthorized disclosure, alteration or destruction of that data would results in little or no risk to Thematic, its customers and its affiliates. Examples of Public data include press releases, educational resources and case studies. While little or no controls are required to protect the confidentiality of Public data, some level of control is required to prevent unauthorized modification or destruction of Public data.

Classification of data should be performed by an appropriate Data Steward. Data Stewards are senior-level employees of Thematic who oversee the lifecycle of one or more sets of Thematic Data.

Data Collections

Data Stewards may wish to assign a single classification to a collection of data that is common in purpose or function. When classifying a collection of data, the most restrictive classification of any of the individual data elements should be used. For example, if a data collection consists of a user's name, email and social security number, the data collection should be classified as Restricted even though the user's name and email may be considered Public information.

Reclassification

On a periodic basis, it is important to reevaluate the classification of Thematic Data to ensure the assigned classification is still appropriate based on changes to legal and contractual obligations as well as changes in the use of the data or its value to Thematic. This evaluation should be conducted by the appropriate Data Steward. Conducting an evaluation on an annual basis is encouraged; however, the Data Steward should determine what frequency is most appropriate based on available resources. If a Data Steward determines that the classification of a certain data set has changed, an analysis of security controls should be performed to determine whether existing controls are consistent with the new classification. If gaps are found in existing security controls, they should be corrected in a timely manner, commensurate with the level of risk presented by the gaps.

Calculating Classification

The goal of information security, as stated in the Information Security Policy, is to protect the confidentiality, integrity and availability of Thematic Data. Data classification reflects the level of impact to Thematic if confidentiality, integrity or availability is compromised.

Unfortunately there is no perfect quantitative system for calculating the classification of a particular data element. In some situations, the appropriate classification may be more obvious, such as when federal laws require Thematic to protect certain types of data (e.g. personally identifiable information). If the appropriate classification is not inherently obvious, consider each security objective using the following table as a guide. It is an excerpt from Federal Information Processing Standards (FIPS) publication 199 published by the National Institute of Standards and Technology, which discusses the categorization of information and information systems.

	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.	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.	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.	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.

As the total potential impact to Thematic increases from Low to High, the classification of data should become more restrictive moving from Public to Restricted. If an appropriate classification is still unclear after considering these points, contact the Information Security Officer for assistance.

Appendix A - Predefined Types of Restricted Information

The Information Security Officer has defined several types of Restricted data based on state and federal regulatory requirements. They're defined as follows:

1. **Authentication Verifier**

An Authentication Verifier is a piece of information that is held in confidence by an individual and used to prove that the person is who they say they are. In some instances, an Authentication Verifier may be shared amongst a small group of individuals. An Authentication Verifier may also be used to prove the identity of a system or service. Examples include, but are not limited to:

- Passwords
- Shared secrets
- Cryptographic private keys

4. Customer Supplied Data

Any data supplied by one of Thematic's customers for analysis. This can include:

- csv files from surveys
- data uploaded through the api

5. Federal Tax Information ("FTI")

FTI is defined as any return, return information or taxpayer return information that is entrusted to Thematic by the Internal Revenue Services. See Internal Revenue Service Publication 1075 Exhibit 2 for more information.

6. Payment Card Information

Payment card information is defined as a credit card number (also referred to as a primary account number or PAN) in combination with one or more of the following data elements:

- Cardholder name
- Service code
- Expiration date
- CVC2, CVV2 or CID value
- PIN or PIN block
- Contents of a credit card’s magnetic stripe

8. Personally Identifiable Information

For the purpose of meeting security breach notification requirements, PII is defined as a person’s first name or first initial and last name in combination with one or more of the following data elements:

- Social security number
- State-issued driver’s license number
- State-issued identification card number
- Financial account number in combination with a security code, access code or password that would permit access to the account
- Medical and/or health insurance information

10. Controlled Technical Information

Controlled Technical Information means any information about the infrastructure and controls employed by Thematic that may give an attacker leverage against our systems.

12. Personal Data from European Union (EU)

The EU’s General Data Protection Regulation (GDPR) defines personal data as any information that can identify a natural person, directly or indirectly, by reference to an identifier including

- Name
- An identification number
- Location data
- An online identifier
- One or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person

Any personal data that is collected from individuals in European Economic Area (EEA) countries is subject to GDPR. For questions, send email to security@getthematic.com

Revision History

Date	Notes
08 Jan 2019	Adopted
10 Feb 2020	Reviewing and ensuring is appropriate
22 Dec 2020	Reviewing and ensuring is appropriate