

# Data Visibility Planning Guide

Thales CipherTrust Data Discovery and Classification



# Contents

<b>3</b>	<b>Introduction</b>
<b>3</b>	<b>Purpose of the document</b>
<b>3</b>	<b>Roles and Responsibilities</b>
<b>4</b>	<b>Data Visibility and Security Flow</b>
4	Define/Update Goals
5	Define/Update Policies
7	Configure Policies
9	Data Discovery
10	Data Classification
11	Data Analysis
12	Remediation
13	Monitoring and Audit
<b>13</b>	<b>More Information for Thales CipherTrust Data Discovery and Classification</b>
<b>14</b>	<b>Thales CipherTrust Data Security Platform</b>
<b>14</b>	<b>Glossary</b>
<b>15</b>	<b>Sources</b>

# Data Visibility Planning Guide



## Introduction

Today, organizations like yours create data at unprecedented rates, all the time, and in various locations including local storage, data lakes, and increasingly in public clouds. To protect data and comply with data protection and privacy requirements, you need visibility into the data you are collecting and storing to determine what data is sensitive and what is not. Thales CipherTrust Data Discovery and Classification enables you to get a clear understanding of what sensitive data you have, where it is located, and its risk of exposure. With complete visibility, you can easily uncover and close your gaps, make better decisions about third-party data sharing and cloud migration, and proactively respond to data privacy and security regulations including GDPR, CCPA, PCI DSS, and HIPAA.

## Purpose of the Document

The main purpose of this “Data Visibility Planning Guide” is to ensure that all the digital assets in your organization have an appropriate classification policy defined for them and to help ensure security teams have visibility of those items that require protection. The classification employed for each asset will limit its use and the implementation of security policies around the asset to reduce the risk of data violations that can compromise data integrity or confidentiality.

In particular, this guide covers the following aspects of defining the data classification strategy:

- Providing guidance on orchestrating the data discovery and classification planning in your organization
- Identifying what sensitive data your organization has, where it is, and how it relates to data privacy regulations
- Understanding and defining data sensitivity levels as part of an impact analysis exercise
- Providing guidance on the remediation process
- Identifying the main roles and responsibilities related with this process

## Roles and Responsibilities

The main stakeholder groups to consider are the security, privacy, compliance, and legal teams. Involving these teams in the analysis enables insight into their various perspectives. Below is guidance for the primary roles to include and should be fine-tuned to align with the unique characteristics of your organization. Each distinct role typically has different core responsibilities with respect to the overall process:

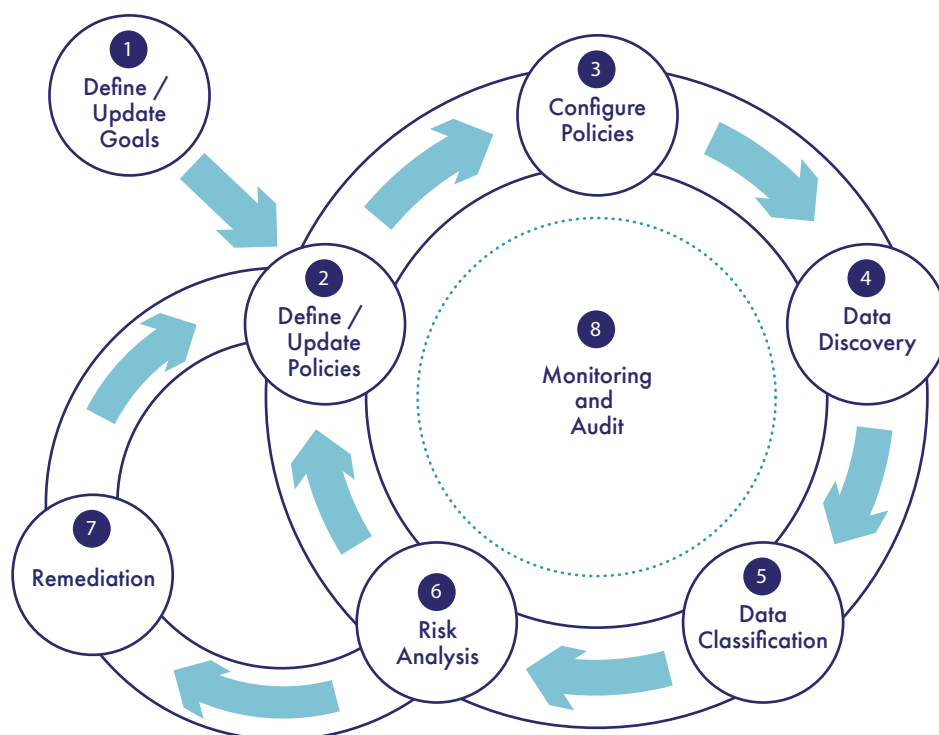
- **Chief Information Security Officer (CISO):** Ensures security across the process, including security of each data location and assisting in the interpretation and application of this process
- **Chief Compliance Officer (CCO):** Aligns the process with laws and regulatory requirements and monitors the process once implemented
- **Data Protection Officer (DPO):** Aligns your organizational data protection goals and monitors the compliance with privacy laws in relation to the protection of sensitive data, ensuring the goals are consistent with the data classification process defined
- **Chief Information Officer (CIO):** Delivers information technology services that meet the requirements of this guide, is involved in the risk analysis, and should provide information for security planning and implementation

- **IT:** Configures the defined policies in the data discovery and classification tool
- **Managers:** Comply with the data visibility and security guidelines, given their scope
- **Data Owners:** Accountable for classifying the data assets they own
- **Data Stewards:** Responsible for classifying the assets of the various data owners (In other words, they help data owners assume their classification responsibilities.)

In a typical organization, the person responsible for the data classification process is the CISO or CIO.

# Data Visibility and Security Flow

This section outlines the recommended data visibility and security flow for a typical organization. The iterative process proposed is made up of eight distinct steps:



**Figure 1.** Data Visibility and Security Flow

## Define/Update Goals

The first step is to define the initial objectives. The overall goal and scope can be updated later to include evolving factors such as new regulations, new functional areas within your organization, or new business requirements.

This step needs to include:

- Definition of the scope of this process, to which groups it applies, and branches and pain points related to data management
- Definition of S.M.A.R.T. goals for the data classification process (where SMART stands for "specific", "measurable", "attainable", "relevant", and "time-bound"); for example, it could be compliance, security, business, or other particular requirements of your organization.
  - Understanding your sensitive data: Visibility into what type of data you have, where it resides, and how it is used.
  - Tracking evolving regulations: Understanding what is defined as sensitive data provides the guidance on how to manage it, how to protect it, and what type of controls are required for regulatory compliance.
  - Reducing data access risks: Map sensitive data and determine which users have access to it, what kind of data it is, what they are doing with it, and then set the scope of remediation programs.

### *Goal and people involved*

The objective is to define/update the goals behind the data classification project.

The people typically involved in this phase are the CISO, CCO, DPO, and CIO.

- Validating digital transformation initiatives: Understand what data could be moved to the cloud, which data requires protection, and what type of protection you need to facilitate your digital transformation.
- Implementing effective remediation: Identify the risk associated with the sensitive data to define the security policies to put in place to apply targeted remediation.
- Identification of the laws and regulations that your organization needs to comply with. Some examples per industry:
  - Data Privacy: GDPR, CCPA, etc.
  - Financial: PCI DSS, SOX, Dodd-Frank Act, etc.
  - Technology/Data Security: FISMA, etc.
  - Healthcare: HIPPA, etc.
- Identification of the main roles involved in the data classification process

## Define/Update Policies

This step implies articulating the strategy to implement the classification process based on the agreed upon goals. Similar to the previous task, this can be modified to reflect the requirements and feedback of the process itself. This step involves analyzing the different types of data your organization manages, defining the required categories and controls to apply, and determining the impact in case of data security violation.

As part of this step, the first task is to analyze the different types of data within your organization. Some examples to review include:

- **Personal data:** name, address, DOB, and email
- **Financial data:** credit card and bank account details
- **Healthcare data:** Medicare card number, European EHIC, and health plan identifier
- **National ID data:** social security number, personal IDs, and tax file number

Next, define the sensitivity levels required for effective management of that type of data. You should consider the following:

- Type of data to manage (such as healthcare data or financial information)
- Type of controls required for each type of data (such as access control, encryption, or tokenization). Users, groups, and processes that should be granted access to sensitive data need to be identified. This will be helpful later in creating the remediation policy to protect sensitive data.
- Type of individuals involved (such as employees, customers, and patients)
- Type of recipients (such as third-party vendors and international third-party vendors)
- Impact of unauthorized exposure of each type of data (such as monetary, business, and reputation implications)

To determine how many sensitivity levels your organization needs, you must have a clear understanding of the expected control granularity. The more control granularity you impose, the more levels you will need to define and manage afterwards, which increases the complexity. If a few categories are defined, some of the data might have more controls than necessary, but the opposite is true as well. So, we recommend a balanced approach that will achieve your digital security needs while avoiding unnecessary operational complexity.

As part of this step you should define the primary loss factors, such as time, business value, and monetary, and the quantitative impacts, such as stock value, fines, professional services cost, business disruption, and any other aspect that your organization needs to measure.

Another factor to address in this step is the impact per data type, such as structured, unstructured, or mail.

The last aspect you need to consider in this step is the roles and responsibilities related to the appropriate classification and use of data assets. In the context of data availability, you have to define who needs to use or access the data to perform the work.

Below, you can find our proposal for four sensitivity levels, already used by Thales CipherTrust Data Discovery and Classification. Your organization may use its own nomenclature for each level or need more or fewer levels, depending on the use case.

### *Goal and people involved*

The objective is to define/update the sensitivity levels, controls to be applied, and loss impact for each level.

The typical involved people in this phase are Managers (as HR and others business unit leaders), CCO, CIO, DPO, and CISO.

Sensitivity Levels Description				
Classification Level Name	Public	Internal	Private	Restricted
Classification Level	1	2	3	4
Impact Type	Lower			Higher
Classification Colour	Green	Yellow	Orange	Red
Description	This type of data can be freely shared with external entities without any harm to your organization.	Low sensitivity. The exposure of this data may not impact your organization, but is not meant for public disclosure.	Disclosure of this data can cause serious damage to your organization. Although less sensitive than restricted, it still requires a high level of protection.	Highly sensitive. Disclosure of this type of data can cause grave damage to your organization, both financially and legally.
Typical examples	<ul style="list-style-type: none"><li>• Price lists</li><li>• Press releases</li><li>• Marketing material</li></ul>	<ul style="list-style-type: none"><li>• Organizational charts</li><li>• Internal newsletters</li><li>• Non-personal data</li></ul>	<ul style="list-style-type: none"><li>• Customer data</li><li>• Employee data</li><li>• Sales information</li><li>• Purchase contracts</li><li>• 3rd party agreements</li></ul>	<ul style="list-style-type: none"><li>• Special categories of personal data</li><li>• Trade secrets</li><li>• Intellectual property</li><li>• Medical formulas</li><li>• Internal financial data</li></ul>
Controls to be applied				
Access Control	None	Internal access within your organization's employees, as required by the scope of work	Access restricted to the team which created the data or collected it	Access restricted to key people who needs to be aware of it, always under protection
Protection	None	Might require some protection	Encryption/ tokenization	Encryption/ tokenization
Per type of data				
Structured	No special treatment	Might require some protection	Tokenize the structured data	
Unstructured	No special treatment	Might require some protection	Encrypt and configure the proper access control to the unstructured data such as .doc, .pdf or .xls files. The segregation of duties could also be improved through the usage of different encryption keys for different groups of users.	
Email	No special treatment	Might require some protection	Only with the appropriate controls in place and include classification level notification. Make sure the recipient is correct before sending. Avoid broadcasting.	
Primary Loss Factor				
Direct monetary loss	None	Define a range of impact depending on your organization's data asset value. The monetary loss could be due to, for example, regulatory laws fines, or customer or employee lawsuits.		
Value of loss of impacted asset	None	Consider the impact on the teams and the time required for restoring the involved asset. For example, if you lose financial data, you probably need to monitor your line of credit, or if customer’s data is compromised, you need to advise them about the breach. It could also impact your organization if IP and trade secrets are compromised		
Value of lost business	None	Consider the business lost in case an internal asset is compromised. Loss of competitive advantage, reputational damage.	Consider the business lost if a private or restricted asset is compromised. This impact could be considerable, for example customer credibility, brand damage, reputational damage.	

## Recommendations

- Identifying all the types of data in your organization can be a big challenge, so organize it according to business processes driven by process owners.
- Start with a small and simple (three or four) sensitivity levels definition, and, based on the analysis done, add more if required. In other words, define the sensitivity levels in a way that is easy for users to work with.
- Regarding the type of controls required per each type of data, take into account each sensitivity level. The definition of who can access and how access is granted should be created.
- A fine-grained policy control enables administrators and system level users to perform their work (such as system backups, updates, and hardware maintenance) without having access to decrypted sensitive information. For defining the access controls for private and restricted data you need to determine:
  - Appropriate access rights
  - Restrictions for specific user roles
  - Restrictions for use sorted by working groups

When access is granted to private and restricted data, segregate access rights management, so you have unique roles assigned to separate individuals.

For instance, allow only finance department members to access critical accounting data, HR to access confidential employee information, and engineering to access development documents. Each Data Store section could be encrypted with an individual key by policy, effectively limiting the access to only those who require it for their work.

- Consider strong, standard-based encryption protocols for data encryption and a high-performance solution for tokenization

## Configure Policies

Once the policies have been defined, use the Thales Data Discovery and Classification solution to automate the data discovery and classification process. First, identify and set up the Data Stores to search for sensitive data (their connection details, physical location, etc.) and second, determine what type of data to find (for example, personal, financial, healthcare data).

Based on the classification policy defined in the previous step, configure the policies in the data discovery and classification tool. The following items should be configured:

- a. Locations to search for data
- b. Type of data to include in the search

To configure the locations to search for data:

1. Catalogue the locations where your data is stored, taking into consideration the data flow. Here are some aspects to check:

- HR data location
- Customer data collection process and storage location
- IP data location
- Mechanism used for exchanging data inside and outside your organization
- Cloud and data lake storage

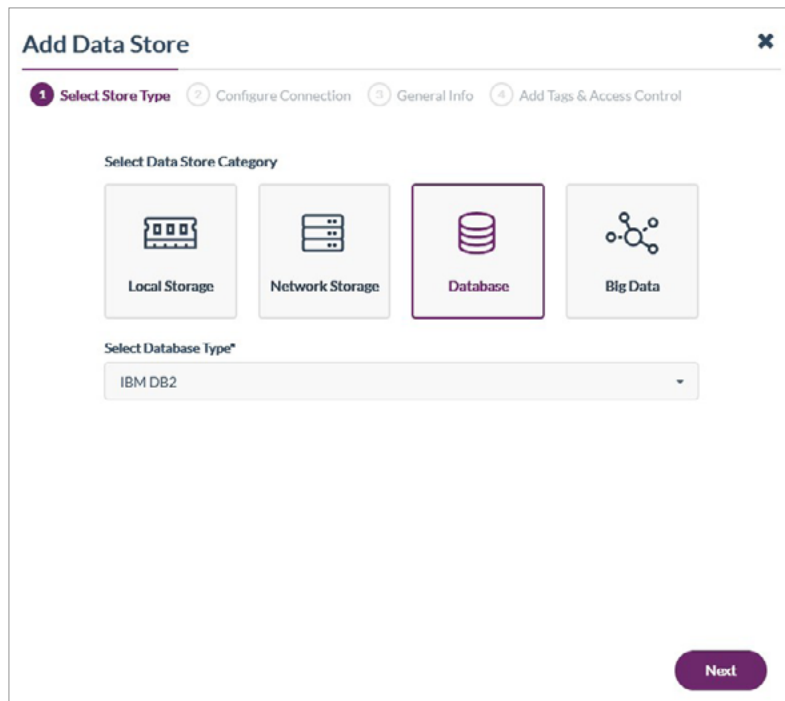
Another important aspect to consider is access to the results of the discovery. This access can be configured as part of the Data Store set up wizard and based on the groups created previously by the user.

### *Goal and people involved*

The objective is to determine which locations and what type of data to include in the search. As part of this step, you have to configure the remediation policies.

The typical people involved in this phase are IT, CCO, and CIO.





**Add Data Store**

1 Select Store Type 2 Configure Connection 3 General Info 4 Add Tags & Access Control

Select Data Store Category

Local Storage

Network Storage

**Database**

Big Data

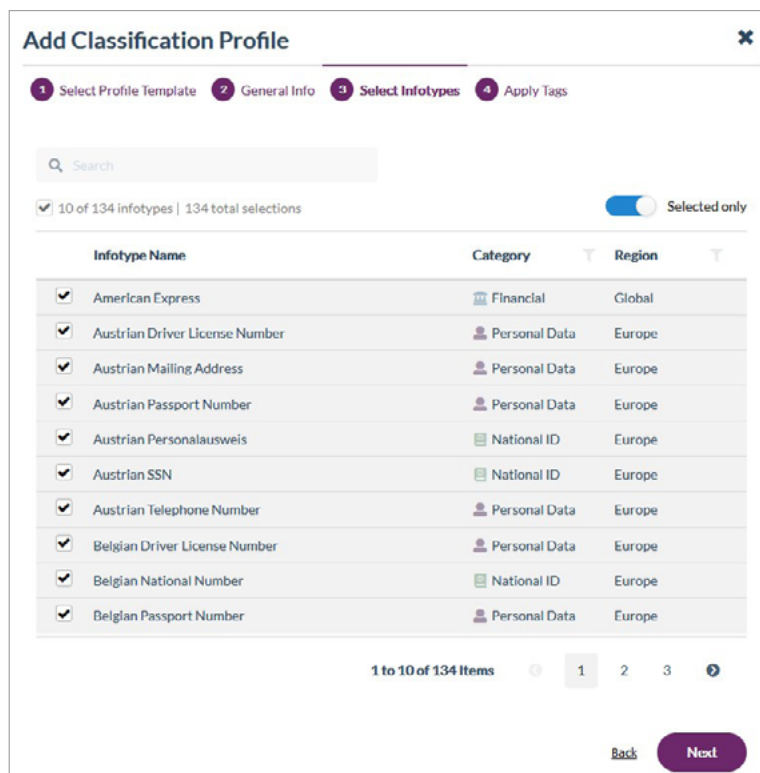
Select Database Type\*

IBM DB2

Next

**Figure 2.** Add Data Store process

2. Define what type of data your organization wants to search for (PII, PHI, etc.), based on the goal that you aim to achieve. Take advantage of built-in data types and pre-built templates; check whether they can be used as-they-are or if they require adjustment to your organization needs.



**Add Classification Profile**

1 Select Profile Template 2 General Info 3 **Select Infotypes** 4 Apply Tags

Search

☒ 10 of 134 infotypes | 134 total selections ☒ Selected only

Infotype Name	Category	Region
<input checked="" type="checkbox"/> American Express	Financial	Global
<input checked="" type="checkbox"/> Austrian Driver License Number	Personal Data	Europe
<input checked="" type="checkbox"/> Austrian Mailing Address	Personal Data	Europe
<input checked="" type="checkbox"/> Austrian Passport Number	Personal Data	Europe
<input checked="" type="checkbox"/> Austrian Personalausweis	National ID	Europe
<input checked="" type="checkbox"/> Austrian SSN	National ID	Europe
<input checked="" type="checkbox"/> Austrian Telephone Number	Personal Data	Europe
<input checked="" type="checkbox"/> Belgian Driver License Number	Personal Data	Europe
<input checked="" type="checkbox"/> Belgian National Number	National ID	Europe
<input checked="" type="checkbox"/> Belgian Passport Number	Personal Data	Europe

1 to 10 of 134 items 1 2 3

Back Next

**Figure 3.** Add Classification Profile process

If your organization manages very specific types of sensitive data, create your own data types and group them based on need.



## Recommendations

- Group the Data Stores, for example, by using tags, so you can analyze the data stores from different angles later on, and so other applications can take advantage of the tags, as well.
- To keep your data safe, plan the access to the results of the scan. For example, set up data access for data owners, CISO, and DPO. Take into account that the groups should be created previously in the Thales CipherTrust Manager, which is central to CipherTrust Data Discovery and Classification, enables your IT organization to set policies, and links to the CipherTrust platform of data protection products to enact remediation.

## Data Discovery

To start the data discovery process, configure a scan using the Thales Data Discovery and Classification solution. As part of the scan configuration, define the Data Stores, the type of data, and when to search.

Data discovery is an iterative process that will provide data visibility to your organization. Being able to schedule a scan and automatically apply remediation actions makes the process repeatable and scalable.

The first step is to define **where** to start the search. The approach at the beginning is to fix possible issues, examine false positives (and possible false negatives), and tune the defined policies. Once this is done, the whole of your organization has to be scanned for sensitive data, considering on-premises, cloud storage, data lakes, and big data. Refer to the “Recommendations” section for more details.

When you already know which Data Stores and data types to scan, select the Classification Profile according to your needs. You can start with a pre-built template or create a custom Classification Profile with a reduced number of InfoTypes to simplify the analysis.

The last aspect to configure is **when** to run the scan. A scan can be executed right after it has been created or scheduled to have data insights updated periodically. This is normally done outside of regular working hours. Our solution provides three options to schedule scans: daily, weekly, or monthly. It is possible to select when to start and end, and the time zone to use.

If you want to run your scans outside of working hours, you can pause the scans during these periods.

### *Goal and people involved*

The objective is to run the discovery process.

The typical involved people in this phase are IT.

## Recommendations

- A suggested gradual strategy for discovery could be the following:
  1. Select a small area to start scanning. Consider starting with frequently used Data Stores, as data stored in those directories is the most exposed, because many people are accessing it. Take advantage of the path definition step during the scan configuration, which will allow you to reduce the scope of the scan to a specific folder. This path provides a starting point that can be easily evaluated for errors (such as connectivity issues, inaccessible locations, or paths with testing data), false positives (and possible false negatives) which will help you tune the policies.
  2. Choose a bigger area to scan for sensitive data, such as a complete local storage or database. Examine the results and fine-tune the policies even more, if required.
  3. Finally, configure several scans to discover sensitive data stored across your whole organization.

This approach will run more scans but the results will be faster and easier to analyze. The information can be grouped for analysis in the reports section.

The first two steps will provide insights about the time required for running scans and how to configure policies and plan accordingly. Some aspects that could affect the time required for each scan are the amount of data to scan, the number of InfoTypes to search for, or the speed of the network, among others.

This path provides a starting point that can be easily evaluated for errors (such as connectivity issues, inaccessible locations, or paths with data to test), false positives (and possible false negatives) which will help you tune the policies.

- Always use the local agent, if available. Even for databases, the agent could be installed locally in the host where the database resides.
- Whenever possible, use specific paths inside the data stores, the more precise the better.
- Take advantage of the built-in templates; consider that there are regulations, such as GDPR, that include a wide variety of sensitive data, and creating templates yourself is time-consuming. Still, make sure the built-in template is useful as it is, otherwise, create your own profile, for example, in this particular case, based on the GDPR template.
- If there are many InfoTypes that you want to search for (e.g. 20 different InfoTypes), consider doing two separate scans to search for ten InfoTypes each. This produces quicker results and increases productivity, as the results of the first scan can be analyzed while the next scan is running.
- Select InfoTypes from a reduced number of regions (one or two) when configuring a scan to reduce the occurrence of false positives.
- Avoid searching for multiple profiles like GDPR, PCI DSS, CCPA data in one search as you will get many matches and it will be difficult to analyze the results. Use one Classification Profile per scan. You can group the scan results in the report for analysis.
- At the beginning, avoid using custom InfoTypes, as this may introduce some error and the results would be difficult to analyze.
- Schedule regular scans for all your Data Stores to be aware of the new sensitive data generated that may add new risks to mitigate and to re-classify data as its value can change over the time.

## Data Classification

This step is key for understanding the type of data that your organization holds. During the discovery phase, the scan engine searches for the type of data requested and, categorizes the data matched by the pre-defined criteria. Once the scan has ended, the results can be analyzed using the different types of reports.

It is possible to group scans in the same report to simplify analysis. For example, if you have scanned different Data Stores for GDPR (using the GDPR Classification Profile), you can group all the scan results in one report and understand how many sensitive Data Objects and occurrences you have, in which type of files they are located, and so on -- all from the GDPR point of view.

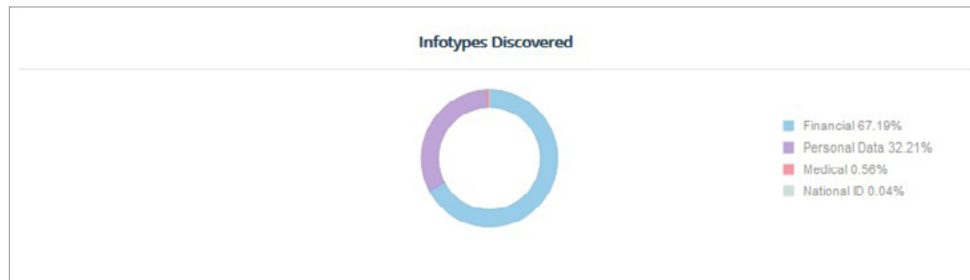
In the Scans tab report, you can visualize the different types of data discovered and the number of occurrences per each type, grouped for all the scans selected in the report.

### *Goal and people involved*

The objective is to categorize data based on pre-defined criteria to manage it properly, understanding whether it is subject to a privacy regulation, how to protect it, and who should have access to it.

The typical involved people in this phase are Data Owners, Data Stewards, CCO, CIO, and DPO.

In the example below, 67.19% of the data in the scanned locations is financial, 32.21% is personal, 0.56% is medical, and 0.04% is National IDs.



**Figure 4.** Pie chart of the InfoTypes discovered.

In the Data Stores tab report, your organization can visualize the sensitivity levels of the different Data Stores scanned, as displayed in the third column of the table included in figure 5.

Scans Data Stores Data Objects									
Search									
Store ↑	Risk	Sens. Level	Scan Name	Last Scan	Location	Infotypes	Data Objects	Sensitive Objects	
▶ Customer Info Ser ...	347	Private	LGDP Audit Scan	22 Jul 2020	São Paulo	13	39	20	
▶ Employee DB	11669	Restricted	LGDP Audit Scan	22 Jul 2020	Rio de Janeiro	12	2	2	
▶ Financial Info Se ...	503	Restricted	LGDP Audit Scan	22 Jul 2020	São Paulo	20	117	65	

**Figure 5.** Table of the Data Stores scanned with their risks and sensitivity levels.

## Recommendations

- Group the scans in the reports by Classification Profiles to understand all sensitive information of a specific set of data types. Alternatively, group the scans by Data Store to understand all the sensitive data in a specific location.

## Data Analysis

The result of the scan provides insights for analyzing the risk of exposure for your organization.

This step is key for planning the best approach for mitigating or removing the existing risk through remediation. Also, it enables you to complete a risk assessment of sensitive data.

Some of the parameters considered for the risk score for the data object are sensitivity level, number of different data types found, number of items matched, and mismatch in sensitivity level between Data Stores and Classification Profile.

The Data Store average risk is calculated as the total sum of risk stemming from all Data Objects on this Data Store divided by the number of sensitive Data Objects. For databases, a table is considered a Data Object.

Start by analyzing the Scans tab report to understand what type of data was discovered (financial, personal, medical, etc.) and in which type of Data Objects (files, databases, etc.). This provides insights about the type of data that you have, and it will allow you, from a simple view, to check if there is something that you were not aware of and that requires immediate action.

Then analyze the Data Stores tab report to understand the risk of exposure per the Data Store included in the scan. This first analysis will provide you with insights on where you should focus first and what types of data are most prevalent in each location.

Finally, utilize the Data Object tab report to analyze in more depth the risk and compare it with the average Data Store risk (located on the Data Store tab report). Based on this, you will be able to determine effective actions to mitigate the risks at a more granular level.

### Goal and people involved

The objective is to understand your organization's risk of exposure in order to prioritize the actions to take.

The people typically involved in this phase are Data Owners, CIO, CCO, DPO, and CISO.

## Recommendations

- Group the scans in the reports by the Classification Profiles to understand the risk of exposure from, for example, a privacy regulation point of view. Alternatively, group the scans by Data Store to understand the risk of exposure in a specific location.

## Remediation

The last step of this iterative process is to take action based on the risk analysis performed. Actions such as data encryption, tokenization, or a proper configuration of data access will reduce your organization's risk of exposure. Thales CipherTrust Data Discovery and Classification is part of the CipherTrust Data Security Platform that provides a variety of different protection techniques that your organization can use based on your needs.

At this stage, you have a clear understanding of your organization's assets and associated risks. This can guide you as you start the remediation process. In the beginning, the policies and findings need to be tested, so our suggestion is to apply the remediation manually.

Let us take an example using encryption for remediation. First, configure the remediation policies based on what type of protection is required for each type of data. Then, follow these steps:

1. **Access Policy:** Configure the access policies to allow appropriate users, groups, and processes to have access to data.
2. **Encryption Policy:** Define/configure the encryption rules to select encryption keys and configuration.
3. **GuardPoint:** Create GuardPoint<sup>1</sup> to ensure a policy is applied on the path to achieve the intended behaviour after remediation.

These steps need to be done once for a path and do not have to be repeated again in the cycle.

Figure 6 shows how the remediation policies are configured from a centralized console, providing your organization with a coherent protection approach across all the data assets that you manage.

### *Goal and people involved*

The objective is to put in place technical security solutions to mitigate or eliminate the risk of data exposure.

The typical involved people in this phase are IT, CDO, DPO, CISO, and CIO.

The screenshot displays a web-based interface for managing policies. On the left is a dark sidebar with navigation links: 'Clients', 'Policies' (selected), 'Policy Elements', and 'Profiles'. The main area is titled 'Policies' and contains a search bar, a selection indicator showing '0 selected' out of '10 results' from '83 Policies', and a '+ Create Policy' button. Below this is a table with columns: 'Name', 'Type', 'Version', and 'Description'. The table lists 18 policies, all of type 'STANDARD' and version '0'. The policy names are 'aa\_policy\_1' through 'aa\_policy\_18'. At the bottom of the table, there is a pagination bar showing '1 to 10 of 83 Policies' and a page number '1'.

Name	Type	Version	Description
aa_policy_1	STANDARD	0	
aa_policy_10	STANDARD	0	
aa_policy_11	STANDARD	0	
aa_policy_12	STANDARD	0	
aa_policy_13	STANDARD	0	
aa_policy_14	STANDARD	0	
aa_policy_15	STANDARD	0	
aa_policy_16	STANDARD	0	
aa_policy_17	STANDARD	0	
aa_policy_18	STANDARD	0	

Figure 6. Remediation policies list.

<sup>1</sup> A GuardPoint is a folder or directory path that CipherTrust protects and controls. Once a policy is selected and applied to a folder, that path is considered a GuardPoint.

Do not forget to include cloud storage in the remediation policies, as there might be unprotected secrets (for example, passwords or encryption keys) that increase business vulnerability.

It is important that this step does not affect your business processes, user tasks, or administrative workflows. Thales solutions are designed to minimize the impact on the overall system performance.

## Recommendations

- Consider applying encryption to private and restricted files that currently have no protection.
- Consider applying tokenization to private and restricted data within database storages that currently have no protection.
- Define granular access controls: who is permitted to access data; which data and data locations are available to them; when the access can be allowed; and what processes are allowed to access plaintext, copy encrypted files, or even view file system metadata. Access controls are available for users as well as groups defined within the solution.
- Start protecting the items that are the most at risk first.

## Monitoring and Audit

In parallel with the other CipherTrust Data Discovery and Classification processes we've discussed so far, we recommend monitoring to make sure the goals are achieved. The last step of the overall process is to run audits to check that controls are in place and are working properly.

As we have noted, data discovery and classification is an iterative process, so it needs to include monitoring cycles to react to the data dynamism and its distribution. As part of this step, consider:

1. Checking the status of each phase, identifying weaknesses in the process that need to be addressed
2. Providing feedback to the process
3. Running audits using an independent and objective vision (This means that a team external to the program area should execute the audit. This could be done by the compliance officer, an internal or audit department, other program managers, or a combination of these roles. It can also be done by people external to your organization.)

This stage will be used to monitor the changes to the data and provide feedback to the process to update the classification, the sensitivity level, and the other important aspects of data to apply effective remediation based on the new findings.

### *Goal and people involved*

The goal of this step is to make sure that the goal defined by the organization is achieved. This goal could be related to data protection, data migration, data privacy, etc.

The typical involved people in this phase are Managers, CISO, and CCO.

## Recommendations

- If during the monitoring you find that a definition in the process is not being used according to the specification, update it accordingly. It is better to have the process defined as it is working and managing the risks that it implies. For example, if a certain item should not be shared outside your organization, but as part of the deployment process that is required by an external customer, it is better to reduce the sensitivity level for that item and manage the risk accordingly.

# More Information for CipherTrust Data Discovery and Classification Solution

We hope that you are well on your way to discovering and defining the data protection strategy that suits your enterprise. Thales CipherTrust Data Discovery and Classification streamlines the workflow from policy configuration, discovery, and classification, to risk analysis and reporting, enabling you to proactively mitigate risks, enforce unified policies, and respond to evolving regulatory challenges. It is available in agentless and/or agent-based deployment modes for reduced complexity and easy scalability.

Follow the links below to quickly get started with Thales CipherTrust Data Discovery and Classification:

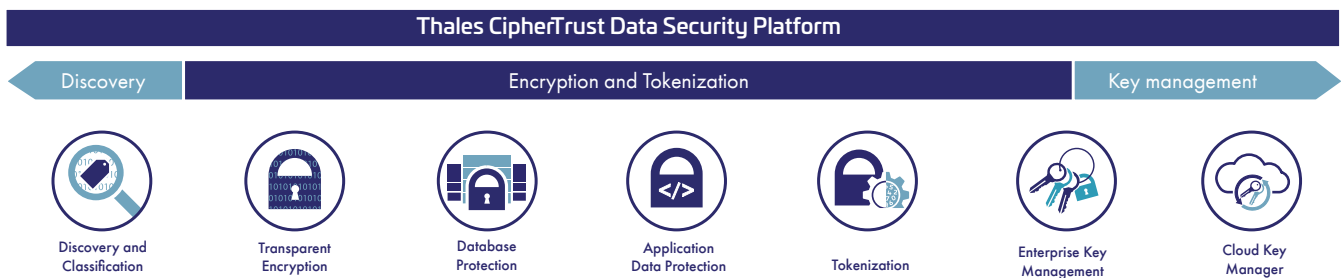
- [Thales Sensitive Data Discovery and Classification](#)
- [Thales CipherTrust Data Discovery and Classification - Product Brief](#)
- [451 Research Highlights Thales's New CipherTrust Data Discovery & Classification](#)
- [Build a strong foundation for data privacy and security - Solution Brief](#)

# Thales CipherTrust Data Security Platform

Thales CipherTrust Data Discovery and Classification is part of the CipherTrust Data Security Platform, providing the following data protection techniques:

- Transparent encryption for files, databases, and containers ([CipherTrust Transparent Encryption](#))
- Application-layer data protection ([CipherTrust Application Data Protection](#))
- Format preserving encryption ([CipherTrust Application Data Protection](#) or [CipherTrust Database Protection](#))
- Tokenization with Dynamic Data Masking ([CipherTrust Tokenization](#))
- Static data masking ([CipherTrust Batch Data Transformation](#))
- Privileged user access controls ([CipherTrust Transparent Encryption](#))

Also, the platform includes centralized enterprise key management (through [CipherTrust Manager](#)).



## Glossary

**Data Discovery:** Process of finding sensitive data in different locations (for example file servers, cloud apps, databases, desktops)

**Data Classification:** Process of categorizing data based on pre-defined criteria – built-in templates, custom types or sensitivity levels

**Remediation:** All actions taken by an organization to mitigate or eliminate the risks

**Data Store:** A place where the data is located, for example, a database, a file server or a cloud service

**InfoType:** Each type of an example for sensitive data, for example, SSN, credit card, and so on

**Data Object:** An element potentially with sensitive data that will be a target for a scan, for example, file, database table, and image

**Classification Profile:** A group of InfoTypes or tuples used together to classify specific Data Objects located in a Data Store

**Sensitivity Level:** A characterization of data based on the evaluation of the potential impact that a loss of availability, confidentiality, or integrity of that data would represent for the organization

# Sources

<https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.199.pdf>

## About Thales

The people you rely on to protect your privacy rely on Thales to protect their data. When it comes to data security, organizations are faced with an increasing number of decisive moments. Whether the moment is building an encryption strategy, moving to the cloud, or meeting compliance mandates, you can rely on Thales to secure your digital transformation.

Decisive technology for decisive moments.





#### Contact us

For all office locations and contact information,  
please visit [cpl.thalesgroup.com/contact-us](https://cpl.thalesgroup.com/contact-us)

> [cpl.thalesgroup.com](https://cpl.thalesgroup.com) <

