devoteam

September 2021

How to identify and **protect your sensitive data**

Creative tech for Better Change

In partnership with

Microsoft

About Devoteam

Devoteam is a leading consulting firm focused on digital strategy, tech platforms and cybersecurity. By combining creativity, tech and data insights, we empower our customers to transform their business and unlock the future.

With 25 years' experience and 8,000 employees across Europe, the Middle East and Africa, Devoteam promotes responsible tech for people and works to create better change.

Creative tech for Better Change

White paper title

Contents

Foreword

Data protection is an essential issue for all organisations and has become the main concern of managers. The volume of data that businesses are dealing with is growing exponentially and remote working makes risk management around protection of information in transit ever more difficult.

Every day colleagues generate documents and share information as part of their ongoing activities, multiplying the possibilities of sensitive data being leaked inadvertently or malevolently. These actions can also generate legal non-compliances leading to substantial penalties.

A simple chat or transfer of files can increase the risk of sharing sensitive data in a manner that is not compliant with the organisation's safety and security policy and the regulations in force.

Here are the most significant questions organisations must face:

- Where is the data stored and who can access it?
- Is the data classified and protected?
- Are data protection policies and practices regularly reviewed and updated?

43% consider as fairly bad their ability to manage data wherever it is located

Source: IDC Compliance Survey for Microsoft - July 2020

Renaud Templier
Devoteam Cybersecurity Director

Introduction

In order to protect data and prevent its disclosure, organisations must respect the standards and regulations in force. To do this, continuity of protection through verifiable controls must be ensured.

> **"Nearly half of companies poorly handle the analysis of exchanged content and the control of internal data leaks."**

40% consider as fairly poor, even very poor their ability to examine data before exchanging it.

Source: IDC Compliance Survey for Microsoft - July 2020

This white paper is based on a specific client case:

> **"How can I identify the sensitive information in my Microsoft 365 tenant[1] so I can comply with my Security policy?"**

AAn international body with more than 100,000 users and several million documents in the Cloud expressed the need to check the maturity level of its data compliance by completing the identification and classification of its sensitive documents in the Microsoft 365 environment.

The border between safety and security issues is not always easy to distinguish but represents a major element in establishing security and compliance policies. Safety is a set of measures to evaluate and advocate appropriate solutions with the aim of protecting persons inside an organisation. Whereas security encompasses measures to protect the resources of an information system.

The objective of this white paper is to respond to the initial question by using Microsoft compliance tools.

[1] **tenant:** Cloud area dedicated to my organisation

From theory to practice

For any organisation, being continually familiar with its confidential information as well as its locations can sometimes be a challenge!

A growing number of organisations have discovered the advantages and efficacy of bringing files together by storing them and sharing them in the centralised environment offered by Teams collaborative solution. Microsoft has invested massively in homogenising and centralising collaborative functions so as to simplify their use and facilitate their adoption by end users. This trend allows organisations to have a global vision of their data protection and compliance.

A practical approach to evaluate the maturity of compliance would be to identify, as much as possible, the level of sensitivity of the data and define the most critical subjects for the organisation inorder to set a search method for the priority documents to be classified.

Meanwhile, the correlation between the strategy and the target governance of the organisation must be taken into consideration.

The real challenges of data protection are:

Correctly associating technologies to reach their objectives

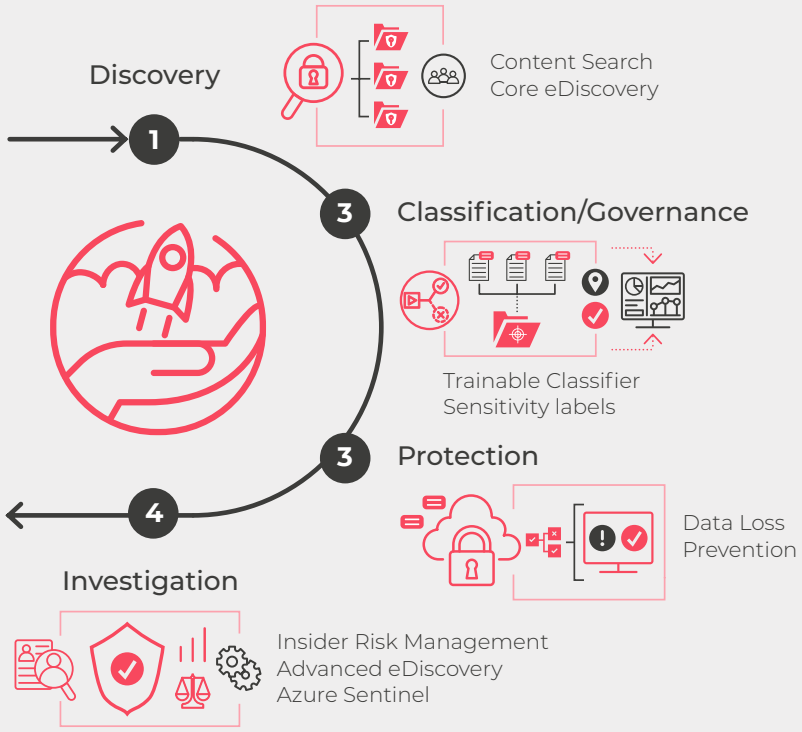
Validating legal authorisation for the right to inspect data

Getting direct accessibility to tools and results

To do this, the Microsoft 365 functionalities and the Cloud Azure services necessary for compliance objectives should be evaluated, depending on the licence level held by the organisation. This evaluation should be regularly refined to benefit from the new features added by Microsoft to its Microsoft 365 offer.

By their level of granularity and their completeness, the features offered by Microsoft allow end to end control of data protection, and thus strengthen the compliance of the organisation's infrastructure.

To carry out this process, the 4 stages to maintain and stabilise are:



1. Discovery

Firstly, a map of the organisation's infrastructure should be created, by studying all the internal documents relating to compliance, as well as the different classification and data protection policies.

This inventory gives a good theoretical vision of what protection level the set of documents present in Microsoft 365 tenant should have. In practice, this phase consists of reassembling a list of documents not aligned with the compliance policy, or with the standards of the organisation.

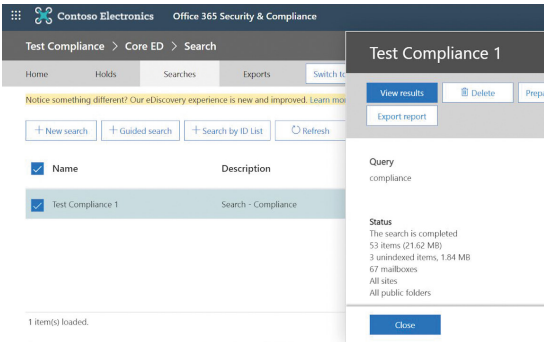
The Microsoft Compliance Centre offers this problem, through two functionalities:

Content search: Allows content in electronic messages, documents and conversations to be found by a single search, through guided or customisable searches.

Core eDiscovery: Offers the same features as the Content Search tool and results can serve as evidence during legal cases. eDiscovery exports are encrypted by security measures and can be analysed via the Advanced eDiscovery tool.

Searches by keywords should be precise to limit searches to the most relevant documents, and it is often essential to fine-tune these searches by iterations, in order to get a usable result (conditions, key words, specific or global locations).

This method creates a key words lexicon which can be reused regularly, and constantly enriched to maintain a good evaluation level of the compliance maturity of the Microsoft 365 tenant.



@Microsoft

2. Classification/Governance

Classifying its documents using labels and governing them with dashboards, is crucial for a company in order to protect themselves. Labels allow traceability and control of information, whether stored or in transit. An exhaustive classification allows analysis of data flows to detect abnormal behaviour and take appropriate corrective measures.

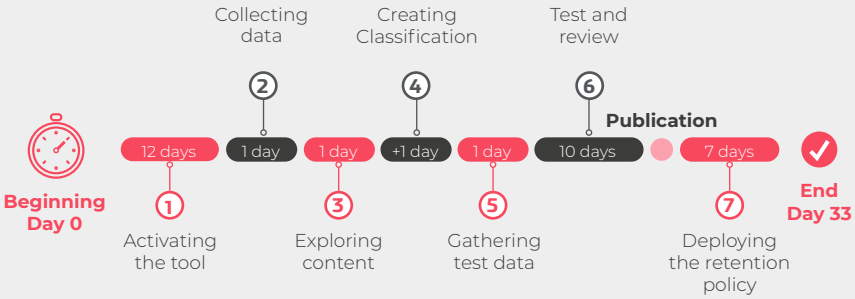
Trainable Classifier

Given the high number of documents processed by each body, to industrialise and systematically process classification, the Trainable Classifier* is a Machine Learning tool which learns to recognise all types of documents: emails, web formats and Microsoft Office suite documents.

To start the process, the organisation provides a sample of several hundred distinct elements so Machine Learning can make forecasts based on the data entered. These results should be confirmed by the organisation to fine-tune their forecasts, sorting false positives and negatives, enabling them to find, classify and label documents appropriately.

* See diagram page 12

Optimal implementation of the Trainable Classifier solution requires a 33 day period following Microsoft's forecasts and recommendations.



@Microsoft * As of late 2020, this solution works in preview and solely with English language elements that are not encrypted.

This solution can also be used at the **"Classification - Governance"** stage as well as at the previous **"Discovery" stage**.

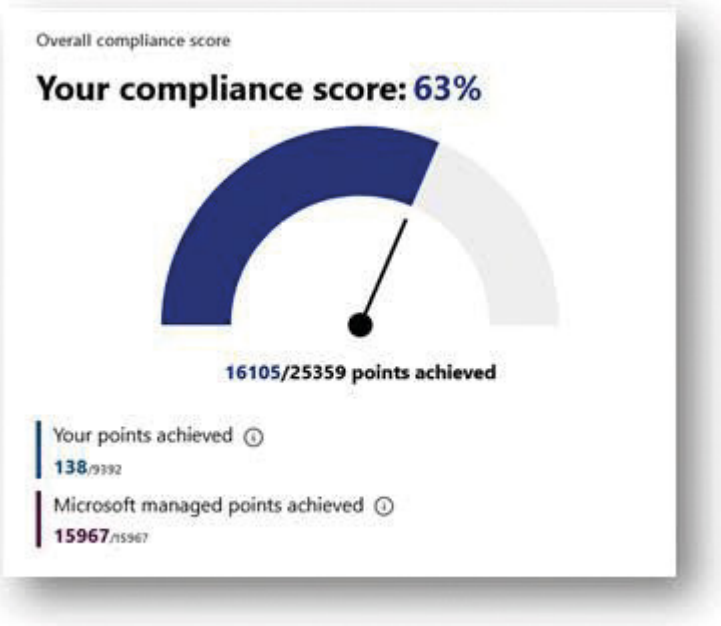
To complement a classification project, an overall view of the organisation's maturity level is possible through the **Compliance Centre** which allows data to be governed globally.

The **Compliance Centre** offers different dashboards identifying and synthesizing information related to the organisation's compliance. These allow visualisation of the global compliance score calculated on the basis of Microsoft's recommendations and suggestions for improvement.

A list of remediation actions that can improve the maturity compliance level is provided and each action improves the global compliance score.

This sensitivity level is determined through the **Sensitive Info Types** function of **Microsoft Information Protection**, allowing recognition of a level of document sensitivity according to data structure, based on regular expressions or a lexicon of keywords (for example identity card, passport or credit card numbers).

The addition of different levels of sensitivity supports the definition of storage, use and communication rules in the M365 tenant.

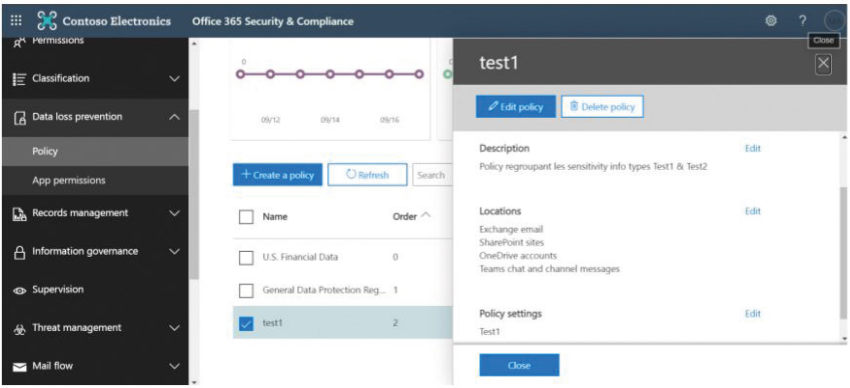


3. Protection

Once non-compliances are identified and documents are correctly classified, the next phase involves protecting information detected as sensitive.

To benefit from labeling and strengthen data protection, Microsoft's Data Loss Prevention (DLP) tool allows written warnings on the strategies base concerning Sensitive Info Types, as well as data protection rules linked to documents or the mailbox according to their criticality.

Stratégie DLP



@Microsoft

These automatically executed rules restrict the circulation of information, applying an encryption policy to them by preventing exposure of unencrypted sensitive data, whether stored, in transit on the network or on the web. The rules should also be integrated into a global debate about data encryption and classification in an organisation's overall information system.

The **Azure Rights Management** service, included in the **Microsoft Information Protection** solution, completes protection by offering encryption, identity and authorisation strategies so sensitive information is seen only by those authorised to access it.

To obtain end-to-end protection against data loss, surveillance and protection functionalities should be extended to the assets of the information system. Microsoft's **Endpoint Data Loss Prevention** (Endpoint DLP) solution detects usage situations and the sharing of sensitive data in the information system to avoid any risky behaviour compromising its integrity.

For Windows 10 managed stations, Microsoft Endpoint DLP offers audit and restriction functions on activities like upload to a cloud service, use of inappropriate browsers, copying to other applications, copying to USB media or network directories, or printing of documents.

4. Investigation

For continued improvement, it is important to investigate internal behaviour contrary to the organisation's policy (illegal, inappropriate, unauthorised and/or unethical behaviour), check legal aspects linked to data and supervise the process of increasing compliance maturity levels.

For this, Microsoft offers **three solutions:**

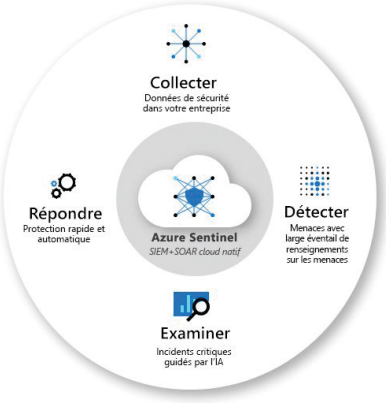
Insider Risk Management identifies risky human behaviour such as theft or leaking of data. This tool allows analysis and the taking of adequate measures to minimise the organisation's internal risks.

To ensure that communications comply with the organisation's standards, the "**Communication Compliance**" functionality offers a thorough analysis of communications in Microsoft 365, relating to the use of inappropriate expressions (inappropriate or shocking language, contrary to the organisation's general communication policy, etc.) both internally and externally. Among the most significant areas of compliance are company strategies, risk management, and regulatory compliance.

Advanced eDiscovery offers a complete solution to the needs of a posteriori investigation, internally and externally, concerning incidents specific to the organisation's data. It also supports legal teams in managing the depositaries involved, as well as the workflow of lawful storage.

Azure Sentinel plays the role of SIEM2 and SOAR3 providing supervision that supports and facilitates the collection of actions carried out by the set of users, items of equipment and applications stored in the Cloud or on-premise.

This solution centralises information which can highlight threats across the whole organisation, providing a global tool for detection, visibility and remediation of non-compliance in relation to the policy defined.



@Microsoft

The Sentinel package allows possible inconsistencies with the organisation's compliance policies to be highlighted, and therfore readjustment of the overall evaluation process of the M365 tenant's compliance, increasing the level of compliance maturity iteratively.

Data protection lifecycle

Compliance Center	A portal simplifying compliance management by calculating a score based on risks, measuring your progress towards the execution of the actions recommended.	Microsoft 365 E3/E5	
Content Search	A search tool in the security and content compliance centre for searching e-mails, documents and conversations in collaborative tools such as Microsoft Teams and Microsoft 365 Groups.	Microsoft 365 E3	
Advanced eDiscovery	The Advanced solution provides an end-to-end workflow to preserve, collect, examine, analyse and export content responding to investigations internal and external to the organisation.	Microsoft 365 E5	
Core eDiscovery	In Microsoft 365, Core eDiscovery provides a basic eDiscovery tool, which organisations can use to search for and export content in Microsoft 365 and Office 365.	Microsoft 365 E3/E5	
Trainable Classifier	Trainable Classifier is a tool using artificial intelligence, which learns to recognise all types of documents: emails, web formats and Microsoft Office suite documents.	Microsoft 365 E5	
Data Loss Prevention	DLP strategies allow automatic identification, monitoring and protection of sensitive data.	Microsoft 365 E3/E5	

Azure Information Protection	Azure Information Protection helps protect sensitive information on Cloud services and on-premise. The AIP tool classifies and labels information based on sensitivity and creates different levels of protection, as well as visual markings.	Microsoft 365 E3 (P1) E5 (P2)
Azure Rights Management	Azure Rights Management is the protection technology used by Azure Information Protection. Azure RMS is a Cloud-based protection service using encryption, identity and authorisation strategies to secure files and e-mails on several devices	
Azure Sentinel	Azure Sentinel is an information and security events management platform (SIEM & SOAR) native to the Cloud which uses integrated artificial intelligence for rapid analysis of large volumes of data inside an organisation.	Microsoft 365 E3/E5
Insider Risk Management	Insider Risk Management allows critical internal risks to be identified and appropriate measures to be taken to reduce them.	Microsoft 365 E5
Communication Compliance	Communication Compliance offers a thorough analysis of internal and external communications in Microsoft 365 for inappropriate language (vulgar, sexist, racist and so on).	
Microsoft Cloup App Security (MCAS)	This offers high visibility, control over data movement and an advanced analytic tool to identify and resolve cyber-threats on your overall Cloud services.	Microsoft 365 E5

Objective: “data protection by design”

By 2025, 175 zetta octets (1012 Go) of data will be present in the global sphere, of which 30% processed in real-time.[4]

“This projection underlines the vital importance of protective strategies for sensitive data, enhanced still more by the evolution of the standards and regulations applied.

Also, the figures mean data protection must be universal and no longer only based on stored data.

Attacks involving data theft will increasingly predominate over the infrastructural attacks which are now the core target.”

[4] IDC for Seagate https://www.seagate.com/fr/fr/our-story/data-age-2025/

Data protection should, like safety, become an integral part of the thinking of all organisations handling sensitive data. The priority for managers should move from a policy based on tools to a global policy of protection of information to achieve “Data protection by Design”, just as we now have “Security by Design”.

In the same way, preventing data violation and unauthorised sharing of documents must involve increased awareness and support for users in understanding the organisation’s data protection tools and policies.

In recent years, the development of the compliance functionalities provided by Microsoft 365 has encouraged the proactive management of data protection processes, helping organisations to meet these new challenges.

“Today searches around “big data” show that data will be the “black gold” of the coming years, and any organisation which cannot properly protect its data will be exposed to serious risk. ”

Vincent Tostain
Devoteam Technical Lead Cybersecurity

From platform to users, the security policy must be all-pervasive

The core business challenges met by our CIOs today mainly converge on a single goal: controlling information throughout its life cycle. Whether it is confidential or public, its mishandling can be a significant risk for the enterprise.

A recent survey carried out by Microsoft among a CISO (Chief Information Security Officer) group reveals changed priorities in a remote working environment:

Offering employees safe access to resources, applications and data is among their top 3 concerns.

Microsoft has invested massively in improving the processes and tools provided to its customers to strengthen the control of information. This white paper presents some of the tools that have come out of this investment such as printed Machine Learning solutions like 'Trainable Classifiers', or protection against data leaks at the workplace with 'Endpoint DLP', or the monitoring of internal risk linked to deviant behaviour with 'Insider Risk Management'. Major advances have been made in dialogue with the outside world, via '3rd Party Data Connectors' or Microsoft's CASB (MCAS), which is now integrated in the Compliance Centre.

But risk reduction is not limited to tools. The complexity of the regulations in each sector is linked to the difficulty of finding internal talent capable of understanding the models and the growing volume of articles for each regulation. Translating these points of control into technical implementation must involve greater collaboration between entities.

Increasingly, we see enhanced communication between human resources and other departments like legal or compliance, acting as relays for an enterprise's lesser-known contact points.

Development of human processes including digital security should obviously respect the user experience. Attention must be paid to change management, as what information should be protected is not always obvious to an employee - even one accustomed to the risks.

This is still truer when personal interaction is linked to the performance of a collaborative environment. If we take the example of Microsoft Teams, the important questions of Governance and security are strongly linked to company productivity. For this, tools are needed, but also best practices should be put in place which simultaneously offer protection, control and visibility - in a benign context.

In short, Microsoft's Compliance solutions can make the difference in this approach to data protection, with the "smart" capacity to combine different signals from heterogeneous environments (multi-cloud, multi-platform, multi-service). Microsoft's laboratories now offer new artificial and cognitive intelligence technologies to transform cyber-safety. If you'd like to know more, follow the keywords NER, Syntex, or Purview directly in Microsoft solution.

The recent increase in hybrid and remote working has shown that understanding and improving levels of security is more important than ever. However secure your information might feel, it is crucial to invest in safety and compliance, both to reduce risk and protect intellectual property.

Charles Tostain
Microsoft EMA - Global Black Belt Advanced Compliance

Conclusion

There is a reason why 91% of enterprises find it difficult to manage data compliance.

Responding to compliance problems requires a high level of cyber-safety and privacy expertise, but also mastery of the tools that should be implemented in response to these issues.

1/10

Despite the focus on security by design, only 1 company in 10 has succeeded in implementing this principle in their safety strategy.

2020 Devoteam x IDC study - A blueprint for security transformation

These compliance issues will grow in complexity in the future, partly as new regulations are added, but also as the amount of data generated increases. So a pragmatic “security & privacy by design” approach is needed, but also a technological one simplifying and automating responses to these compliance problems to the maximum.

Devoteam is able to respond pragmatically to complex compliance issues by combining its expertise in cyber-security with that of Microsoft security add-ons, notably in relation to compliance.

What can be done operationally and where do we begin?

1 Classification and data protection design policies corresponding to operational reality.

2 Adopt a pragmatic approach driven by the risks

- Define use cases — *"Working in close collaboration with core businesses to have concrete cases"*
- Determine the risks linked to these use cases — *"Defining the most frequent cases and those having the most impact because compliance was not observed"*
- Prioritise them according to their frequency — *"An analysis of risks to determine the criticality of use cases and fine-tune the prioritisation of scenarios to consider"*

3 Proceed by iteration based on our methodology as developed above:

Discovery

Classification/ Governance

Protection

Investigation

Karim Bouami
Devoteam Digital Identity & Trust & Services Director

About Devoteam M Cloud

With 500+ clients, Devoteam M Cloud is one of the world's leading providers of Microsoft Cloud technologies with currently 16 gold certifications and 4 Advanced Specializations - Kubernetes on Microsoft Azure, Adoption and Change Management, Modernization of Web Applications to Microsoft Azure, Windows Server and SQL Migration to Microsoft Azure. Our 800+ Microsoft Experts in EMEA offer to medium-sized companies and large enterprises a solution and product portfolio that enables large portions of digitization, new forms of collaboration and makes in-depth analysis of company and production data a reality. Devoteam M Cloud modernize your entire IT architecture, accompany our customers on their journey to the cloud and make it fit for the digital future.

Key figures

500+ clients

1300+ certifications
and 800+ Microsoft Experts in EMEA

16 gold certifications
and 2 Advanced Specializations

Microsoft Gold Partner

Microsoft | Multi Area Partner Orchestration

Microsoft in Education Global Training Partner
Microsoft

devoteam
M Cloud

devoteam.com/partner/microsoft

The authors & contributors

Vincent Tostain
Devoteam Technical Lead Cybersecurity
vincent.tostain@devoteam.com

Karim Bouami
Devoteam Digital Identity & Trust Services Director
karim.bouami@devoteam.com

Ludovic Hecky
Devoteam Cybersecurity Senior Consultant
ludovic.hecky@devoteam.com

Amélie Barboteau
Devoteam Cybersecurity Senior Consultant
amelie.barboteau@devoteam.com

Charles Tostain
Global Black Belt Advanced Compliance
charles.tostain@devoteam.com

With the contribution of Devoteam's and Microsoft's Marketing teams.
Contact: Marketing team marketing-fr@devoteam.com


Creative tech for Better Change