

Data & More | Data Minimization Manager™

Remove privacy data from mail, OneDrive, SharePoint, and FileShares to protect digital privacy and minimize the risk, cost, and liability of processing unstructured data.

01 Executive Summary

In the end, all successful data privacy initiatives end with **data minimization**. That is, deleting data that is illegal to keep under data privacy laws such as GDPR, CCPA, and PIPEDA. Now, deleting data might sound easy—and it is. However, only deleting the correct data is a whole different story. To delete **illegal data**, you must first classify the data and identify the data subject, then verify your automated findings with the data owners, and only then can you **delete the illegal data**.

Luckily, we have been working on just that for the last seven years, both to implement and operate data minimization at scale, from organizations with fewer than 40 employees to those with over 40,000. Using the same simple approach of classify | verify | minimize. Building on this straightforward concept, our platform offers a comprehensive solution for data minimization.

If you want a demo of Data & More, you can try demo.dataandmore.com, and we will sponsor up to 25 test users for up to 2 months.

w: dataandmore.com // m: support@dataandmore.com

Table of Contents

01 Executive Summary.....	1
02 About Data & More.....	2
03 What is the Data & More Privacy Platform.....	3
04 Why do we need Data Minimization.....	3
05 The Data Minimization Process.....	4
05 Data Classification.....	5
06 End-User Verification.....	9
07 Data Minimization.....	12
08 Compliance Analytics.....	13
09 Data & More Purview Integration (add-on).....	15
10 Data & More Copilot Privacy Protection (add-on).....	16
11 Get a Sponsored PoC.....	18

02 About Data & More

Data & More was founded in 2016 by a European team of Data Scientists and Data Privacy advocates. When the EU rolled out GDPR in 2017, the Data & More | Privacy Platform quickly became the de facto standard for [Document Classification](#), [Data Breach Mitigation](#), and [Data Minimization](#). Over the last few years, the Data & More | Privacy Platform has evolved to a full GRC suite for unstructured data. In 2024, the platform is used by more than 150,000 users in over 28 countries, including the US and Canada.

We provide our services to private, public, and non-governmental organizations, including law enforcement and defense. On a daily basis, we monitor over 3.000.000.000 documents in an effort to protect hundreds of thousands of data subjects.

Data & More is [ISA 3000](#) and [ISAE 3402](#) certified and 100% owned by the people working in Data & More.



03 What is the Data & More Privacy Platform

The Data & More | Privacy Platform is a comprehensive tool that enables organizations to identify, verify, and minimize privacy data in large unstructured data sources such as mail, file share, OneDrive, and SharePoint. It has severely award-winning features that have made it the platform of choice for CIOs, CISOS, and DPOs.

- Comprehensive International Privacy Data Classification (29 languages)
- Multiple Data Privacy Framework such as GDPR, CCPAM IAAP, and PIPEDA
- Data Subject Identification and Data Subject Access Request
- Enduser Verification that moves the minimization responsibility from IT to the data owner
- Build-In Data Minimization
- High Speed and Scalability
- Privacy by design
- add-in for Microsoft Copilot Protection
- add-in for Microsoft Purview
- add-on for Data Breach Mitigation

04 Why do we need Data Minimization

The simple answer is that 137 out of 194 countries require it by law, according to the UN Trade and Development Program (UNCTAD). In the EU, that is implemented as GDPR; in the US and Canada, different states have different implementations—all requiring data minimization.

From an organizational perspective, it is, first of all, a matter of maintaining trust from the consumers, employees, and shareholders. Suppose an organization loses or misuses personal data. In that case, it can be very costly and, in some cases, lead to bankruptcy, induced by consumer boycotts or by blackmail and extortion from bad actors.

Most organizations can survive a data breach if they only have the data that they are legally entitled to. Still, if they have old privacy data such as old passports, customer records, CVs, health information, and travel books lying around - the cost of a data breach can be insurmountable.

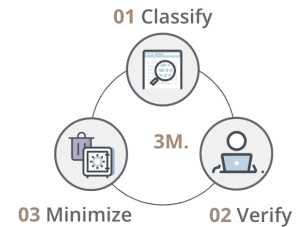
The average cost of a data breach is 4,4 million euros in 2023, and the average price of a fine for non-compliance to GDPR is 1,7 million € in the EU.

Finally, we need to minimize data because it is the right thing to do. Victims of Data breaches are subject to blackmail, having their financial lives destroyed. If we look at the new appetite for AI the risk of having old privacy data lying around in MS 365 or fileshares is increasing vertically. Data Breaches caused by tools like CoPilot are reported daily, sometimes due to bad luck and others due to evil intentions. (For more information, Data Breaches caused by AI [Data & More CoPilot Privacy Protections](#))



05 The Data Minimization Process

To effectively implement a data minimization process, it needs to be straightforward... exceptionally straightforward. If not, it becomes ineffective, leading to errors and vulnerabilities. One of the critical insights to design the D&M minimization process is the probability that a document that is older than three months **AND**, contains Personal Data **AND** is located in MS 365 **AND** will ever be opened again is less than 1 in 100,000. Therefore, we establish a rule called three by 3, which covers the data's 9-month lifespan from creation to deletion.



Classify | Verify | Minimize

In the first three months of a document's lifetime, the data owner can work with the data doing what ever the data is intended for. At the same time D&M Privacy Platform is classifying the data so it knows if the document that needs to be minimized sooner or later. We call the period the "Grace" period, and it is used to read and classify documents. (refer to section 04 of Insights into Classification)

Classify | Verify | Minimize

After three months, the document is added to the End-user Verification Report (see section 05) - and an email is sent to each of the data owners with a link to their reports. In this report, the user can mark any documents as private or for dispensation if they would like to keep them for a longer time. The user has three months to verify the classification and mark data that should not be deleted.

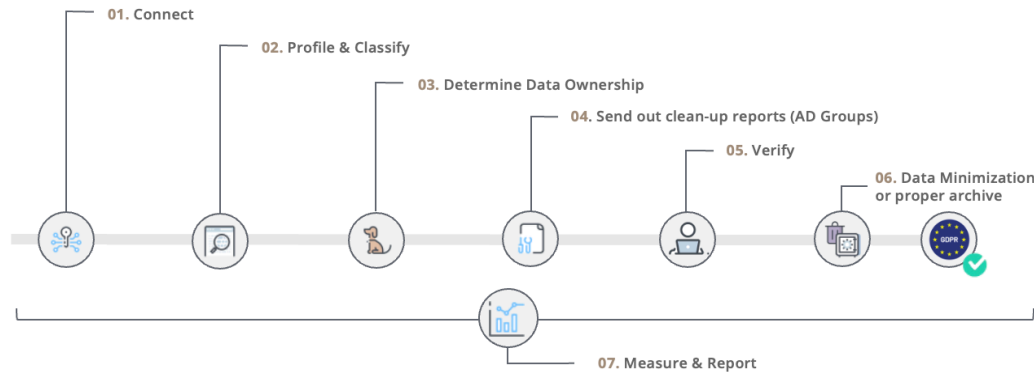
Classify | Verify | Minimize

Now, between six and nine months after the Privacy data has been created (or received), the data will be deleted by the DMCS. If the data resides in Microsoft 365, it can be retrieved within 30 days, after which point only organizations with custom MS 365 backup can recreate the data.

Data residing in file shares will be deleted immediately and can only be restored via backup. Note that if data is restored, it will be deleted again within 6-9 months if no data owner intervenes and marks it for dispensation.



Where to start



Step 1. Data is connected to the DMCS using one of the many intelligent connectors in the DMCS.

Step 2. D&M Privacy Classification is applied to the data, and the standard 3X3 minimization policy is adapted

Step 3. Privacy Data with no apparent owners are identified, and custodians are assigned using the custodian mapping process.

Step 4. The Validation Report is sent out to the data owners and custodians

Step 5 The result of the first classification is verified, and the classification is updated.

Step 6. Data is minimized, aka deleted/moved/encrypted

Step 7. Online Reporting is available for PM, DPO, and Compliance Documentation

05 Data Classification

Data & More's data classification team has spent over six years developing a comprehensive privacy compliance classification system that is used, tested, and quality assured on billions of data each day. This system can identify privacy data as required by law for each country. It is maintained and optimized daily by a multilingual team.

DMCS provides an entire Data Privacy Classification in 28 languages, including local health information, unions, religious and sexual orientation, and all the sensitivity categories required in, e.g., the EU, North America, and Canada. Examples below:



Confidential Personal Data Document Classes (examples)

- European & international ID
- ID card, number, or information
- European & International Social Security info
- Social security card, number, or information
- European & international health cards
- Health card, number, or information
- European & international drivers' licenses
- The card, number, or information
- European & international passports
- The passport, number, or information
- Credit cards
- The credit card, number, or information
- Tax information
- Tax returns, etc.
- Residence permit
- Permits and or information in them
- Salary information
- Pay slips, etc.
- Employment documents
- Contracts etc.
- Recruitment (Application/job offer/CV)
- A wide range of information related to the recruitment process, job applications, CVs, job interviews, etc.
- Bonus agreements
- Dismissal or resignation
- Terminations or resignations
- Written warnings
- Expulsions

Criminal offenses document classes (examples)

- Criminal record
- Criminal records and information about them
- Offenses, fines, and convictions
- Convictions, fines, etc.

Sensitive personal data document classes (examples)

- Health info
- Diagnose



- Illnesses
- Medication
- Sick leave
- Health evaluation
- Prescriptions

Trade union membership (examples)

- Membership of a trade union

Orientations Belief & Origin (examples)

- Which country do you come from
- Ethnicity
- Membership in a political party
- Religious orientation
- Member of a religious church
- Religious congregation
- Gender types
- Information about sexual orientation

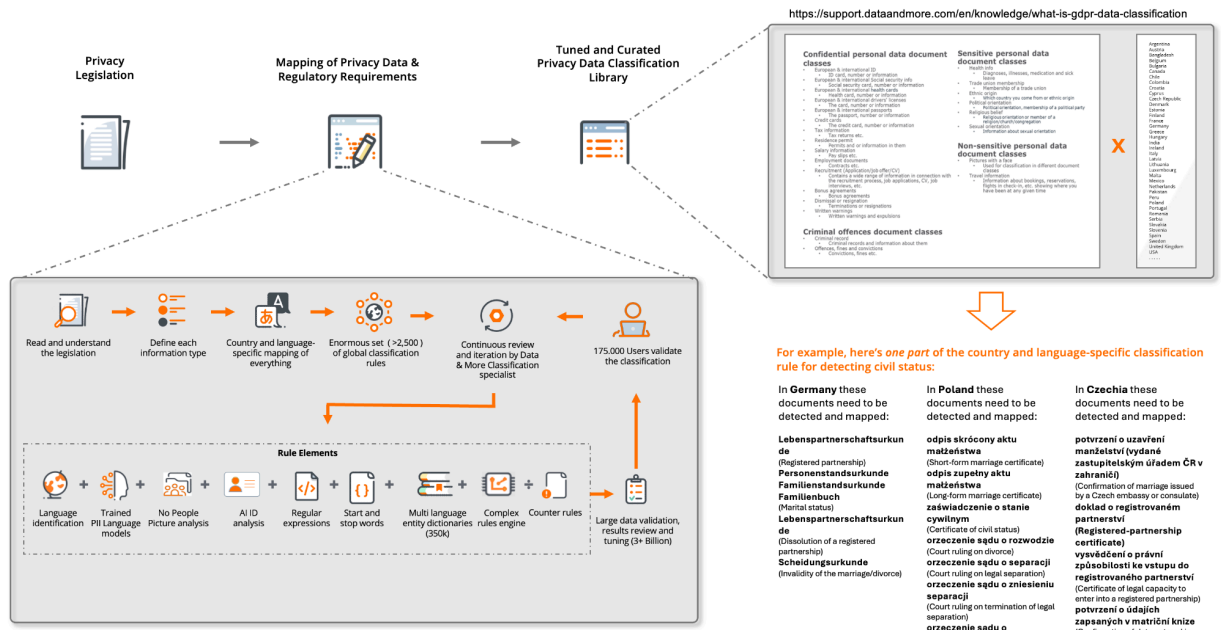
Non-sensitive personal data document classes (examples)

- Pictures with a face
- Used for classification in different document classes
- Travel information
- Travel bookings
- Reservations
- Check-ins, a.g., showing where you have been at any given time

The above list is an example of some of the Privacy classifications. For each class, there are hundreds of thousands of classification elements that are used to identify positives and remove false positives.

Now the implementation of our classification is our **“Secret Source”** and requires a lot of technology and a lot of language expertise. See the illustration below.





(Data & Mores Serect Souce in the form of technology and language integration)



06 End-User Verification

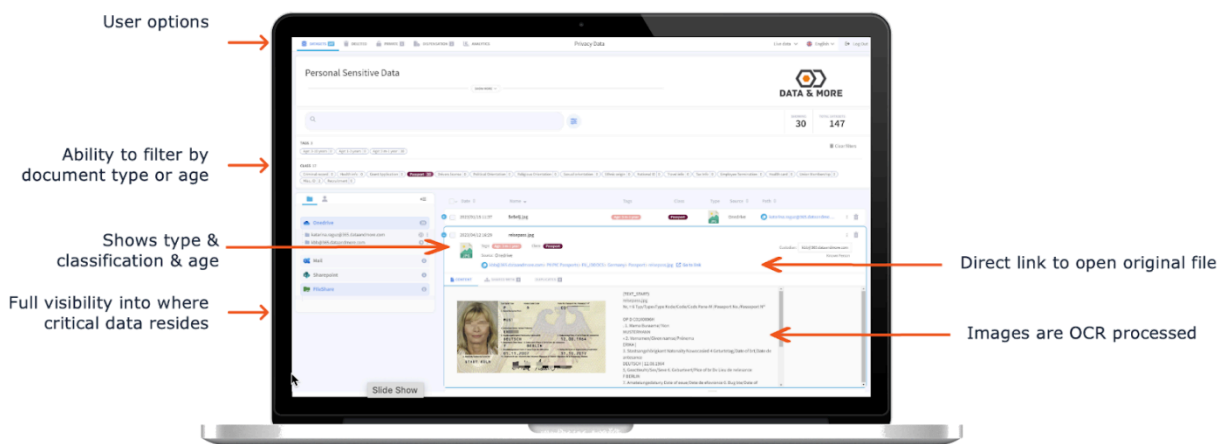
The Verification Report

Data & More End-user Verification gives the data owners a complete overview of the data that has been auto-classified as non-compliant. The data owner can then use the report to mark documents as misclassified, private, or dispensation, depending on the individual's needs.

It gives the ensure quick preview of the and insight into:

- The total amount of data that must be cleaned up
- Document Age
- Document Class
- Document Location
- Document Type
- Document Owen
- Data Subjects

By combining the age filter with location and class, the end-user can very quickly identify, e.g., passports older than five years that are located in the inbox or health certificates older than ten years in an HR SharePoint site.

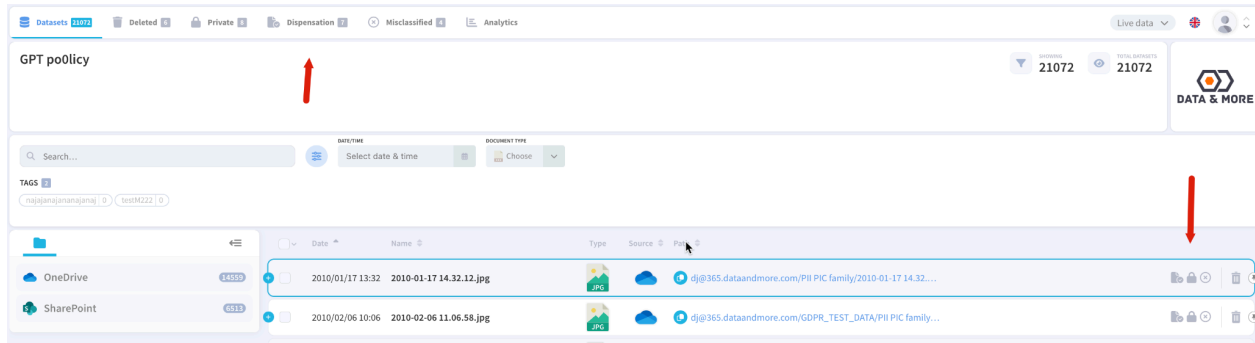


(The End-user Validation Report gives each end-user easy access to their own Privacy Data)

One of the many advantages of the D&M for Purview End User Verification Report is that the end user can quickly check the data and mark data that they need to keep with "Dispensation" or, in the case of Private data, as Private Data. This type of marking can be customized to specific organizational needs.

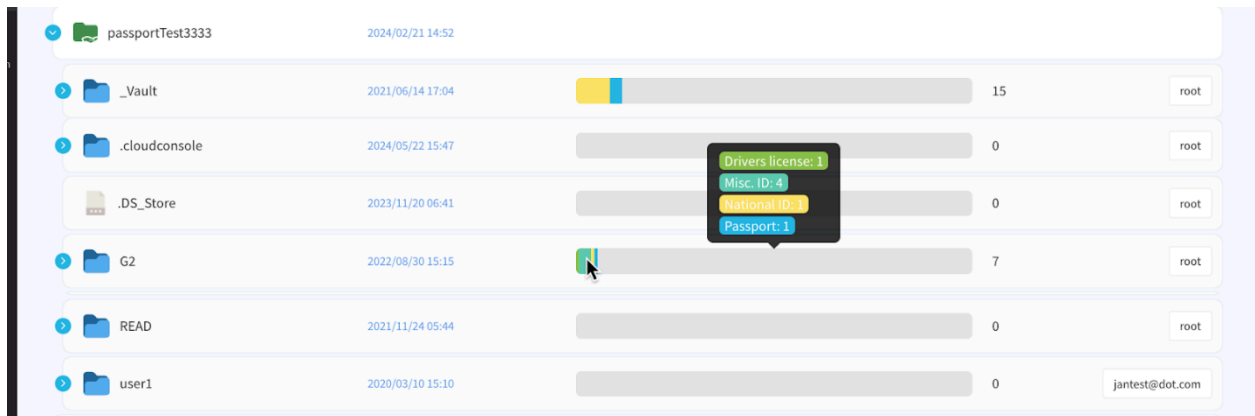


If the end-user thinks that a document has been falsely classified, the end-user can mark it as misclassified as input to the Data & More's classification team.



Custodian Mapping

Proper data management requires that one (or more) person be responsible for a specific data repository. This is easy for Mail and OneDrive, as it is the owner's responsibility. However, with SharePoint and FileShare, this is often much more blurry due to the rotation and churn of employees. Data & More offers a great tool to identify, manage, and maintain "Custodians." The best thing about custodian mapping is that it works by a risk-first approach, so you can start mapping the data that is a risk and leave the rest for later.



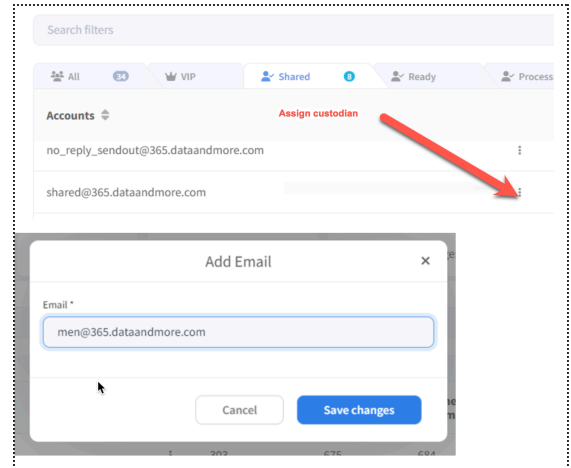
When the custodian mapping is done, all data can be verified appropriately - using the Validation Report.



Shared Mailboxes

Establishing ownership of shared mailboxes is another critical requirement for efficient data minimization. This can be done in multiple ways.

1. If a shared mailbox has an assigned manager (like a normal user), the manager will automatically get custody of the shared mailbox.
2. If a normal user is assigned via the Azure portal as “another email,” that user will get custody of the shared mailbox.
3. Finally, if neither a manager nor another user has been assigned, the DMCS supports manual assignment of the shared mailbox to specific users.



When the custodian has been assigned to the shared emails, all data can be verified appropriately - using the Validation Report.



07 Data Minimization

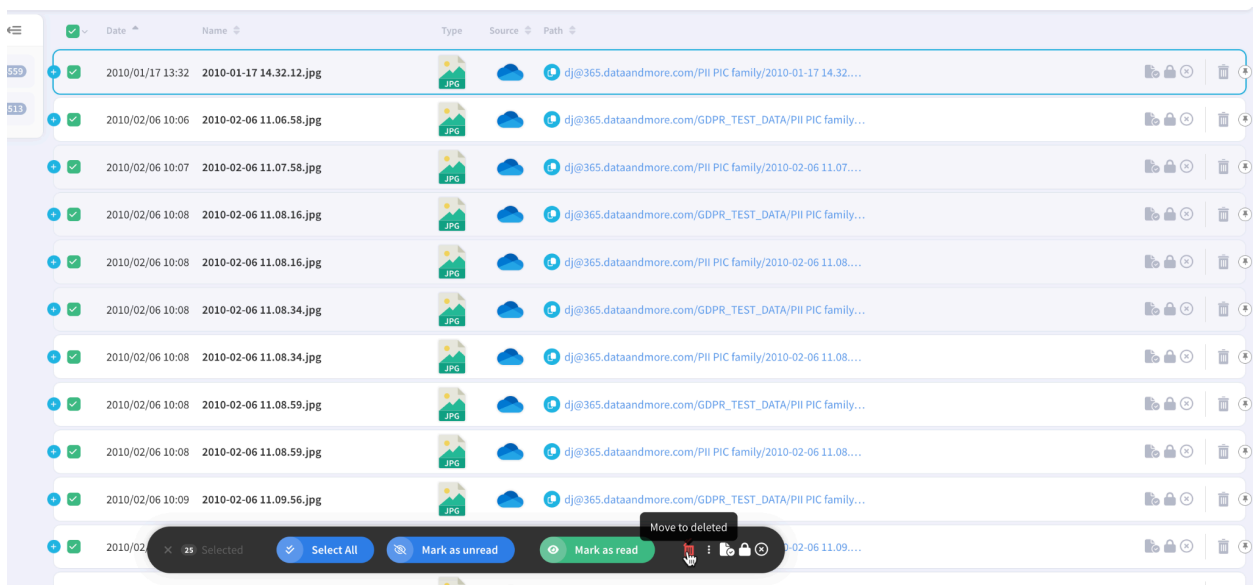
One of the significant advantages of using the Data & More Data Minimization tool is that the responsibility for deletion is moved from the IT department to the owner or custodian of the Privacy Data. The actual data can be configured either som opt-in or opt-out, also known as **Confirmed Delete** or **Automatic Delete**.

Automatic Delete

Automatic Delete, as the wording implies ... deletes the data that has been sent to the data owners for verification after a predefined time period. We recommend deleting data that has been notified after three months if the organization's verification shields follow the recommendation UNLESS a dataset has been marked as private / Dispensation or Misclassified.

Confirmed Delete (optional)

Some organizations prefer to let the data owner or custodian “press the button” for each data set that should be deleted. Now, this might sound more tedious than it is - because a user can select multiple data sets by classification, type, or age and delete them all at the same time.

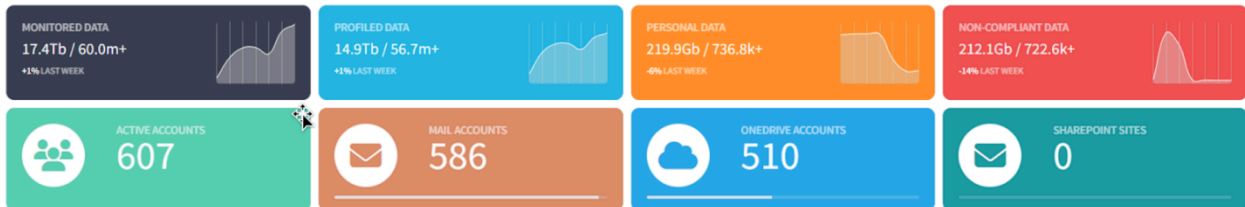


However, experiences with more than 150,000 users show that now all are getting around to clean deleting illegal privacy data - and that is why all our organizations sooner or later turn to **automatic delete**.



08 Compliance Analytics

Compliance analytics serves a dual purpose. First, it helps C-level executives, DPOs, and middle managers ensure that data minimization is proceeding as planned and allows them to follow up on any irregularities in the Data Privacy landscape. This could include data sources with a large amount of unmanaged Privacy Data or individual users who have marked large data sets with Dispensation or as Private. The second purpose of Compliance Analytics is to provide comprehensive and trustworthy documentation to auditors—both internal and external Data Privacy Authorities.



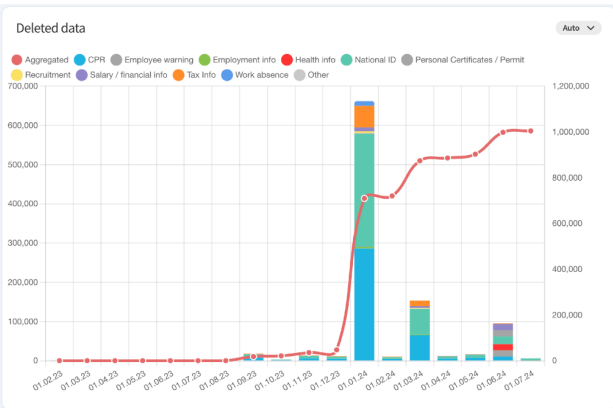
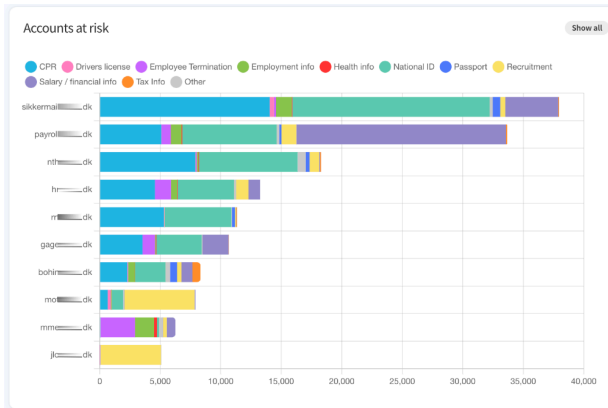
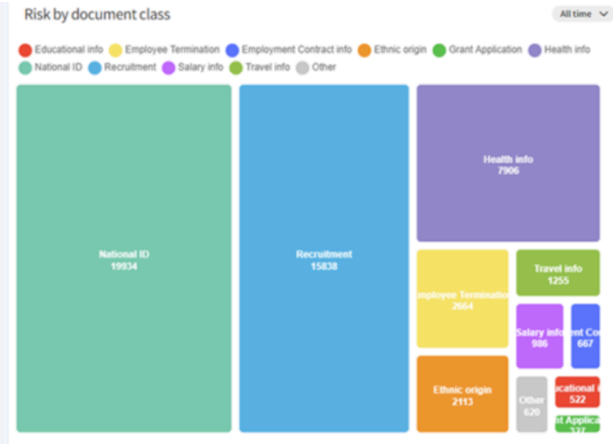
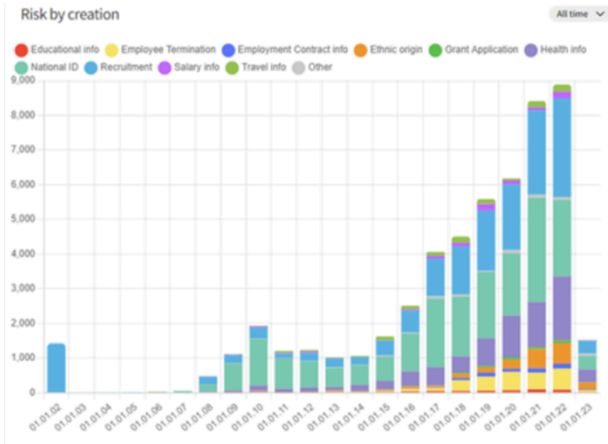
(Dashboard header in Compliance Analytics)

Key metrics in the Compliance analysis include:

- Total number of Data Subjects
- Total number of datasets with Privacy Data
- Document age
- Data location
- Owner or custodian
- Document classes
- Data for data minimization (Owner, Type, Location)



Examples of Compliance Analytics Graphs:



In Compliance Analytics, the report can be filtered by AD Groups or specific policies or tags. The report can be used to show particular document classes or company-wide but will never show the actual content data.



09 Data & More | Purview Integration (add-on)

Microsoft Purview is an excellent tool for MS 365 compliance, Including Data Loss Prevention and in-document tagging. However, more is needed if you want to be compliant with any Data Protection legislation. Let's take a look at what D&M Purview integrated adds to Purview.

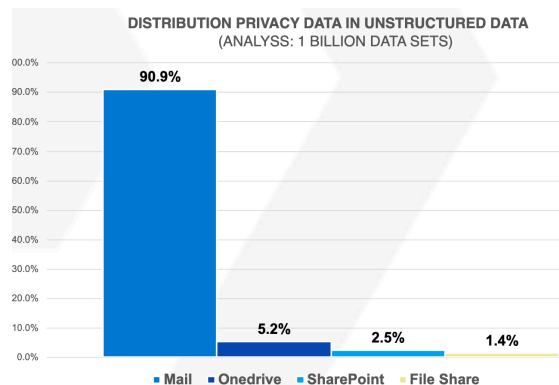
1

1. Purview can protect less than 10% of Privacy Data because it's located in old mail or nonaccessible file types. ([See data and more for Purview for a complete analysis](#))

2. With Data & More Purview integration, you will be able to identify and minimize +99.99% of any privacy data in Microsoft 365

3. Purview gives you the option to build a Privacy Classification based on some preexisting libraries and regex. However, creating an entire Privacy Classification in Purview has never been done and would take a dedicated team multiple years (if at all possible). Luckily, the D&M Privacy Classification can be transferred to Purview via Sensitivity Labels, which enables the full potential of Privacy Protection in Purview, including DLP, Oversharing monitoring, and deletion of legacy data in emails.

4. Finally, by using the Data & More Verification Report, data owners will be able to see what data that have been classified as up for deletion. No matter if data have been classified by Purview automatically, manually by an end user, or by the Data & More Classification Engine. When a data owner can verify that data can be deleted using the Verification Report, the responsibility for deletion is moved from the IT department to where it belongs.



From a Privacy Perspective, Data & More gives the Purview “Super Powers” and unlock a suite of functionality that are based on correct classification and validation, like DLP, sharing control, and Retention policies.

¹ Analyses of 1 billion datasets conducted by D&M 2024



10 Data & More | Copilot Privacy Protection (add-on)

If you subscribe to the Data & More Privacy Platform, you can add Copilot Privacy protection. In short, **D&M | CPP** prevents Microsoft Copilot from accessing Personal Sensitive Data by adding sensitivity labels to the data, with the above limited as outlined in section 08.

The high-level methodology developed by Data & More for ensuring Copilot data readiness consists of the following steps:

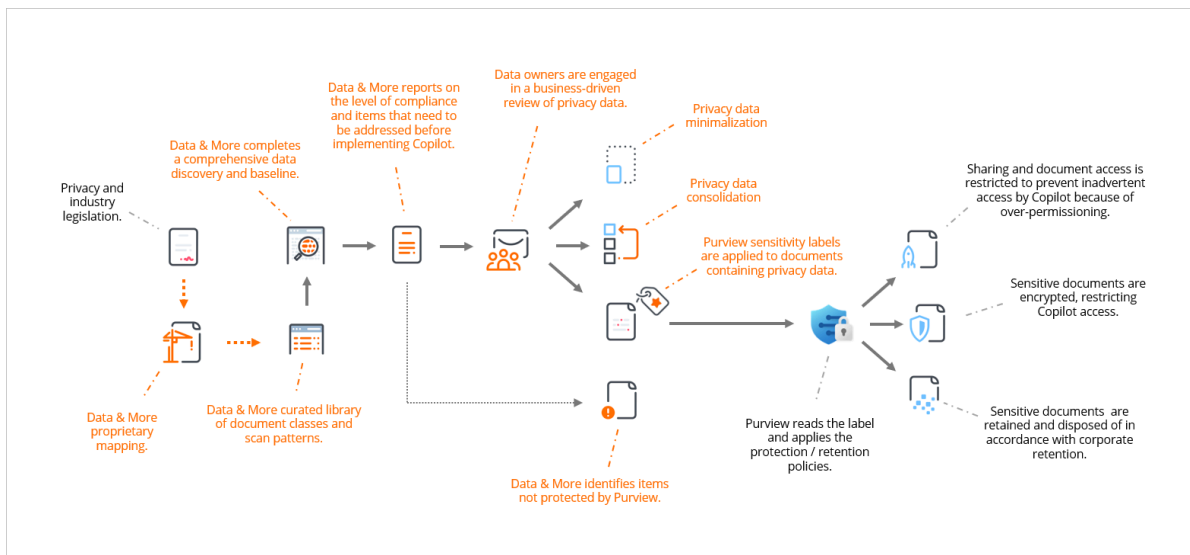
- 1. Deploy D&M | CPP and complete a data discovery.** As mentioned in this architecture section, this requires the creation of an App Registration in Azure. It also requires a location to host the software. This can be the Data & More cloud or a server, physical or virtual, that you host.
- 2. Identify items to be remediated.** **D&M | CPP** will provide a summary of all documents containing privacy data, the type of privacy data identified, the location of each document and its age. It will also provide additional insights, like the occurrence of duplicates and the level of access. This information can be curated into a **Copilot Data Readiness Baseline** and used to determine the best option for addressing each set of privacy data.
- 3. Conduct a business-driven privacy data review.** Data owners and custodians are the best people in an organization to make decisions about privacy data. **D&M | CPP** automates the engagement of these individuals, distributing information about documents containing privacy data to each data owner and custodian, allowing them to mark documents for deletion, retention, relocation, and labeling.
- 4. Configure Purview labels and data protection policies.** The labels needed for encryption and content analysis blocking can be configured along with the policies to implement these services on labeled content.
- 5. Remediate identified data.** Once reviewed by business owners, **D&M | CPP** can:
 - Purge items marked for deletion,
 - Transfer items marked for relocation to the appropriate approved storage location and
 - Apply Purview sensitivity labels to documents so content analysis blocking and file-level encryption can be applied to those documents.

It may also be necessary to adjust permissions and content sharing should any 'gaps' be identified as part of the remediation.



6. **Enable Content Analysis Blocking and Restricted SharePoint Search.** This establishes the scope of Copilot's access to data in alignment with organizational content management practices and data governance, both of which need to support the organization's legal and regulatory requirements.
7. **Complete a Copilot Data Readiness Review and certify the environment as Copilot Ready.** This step reviews the results of the remediation and configuration against the **Copilot Data Readiness Baseline** to ensure all privacy data has been properly addressed. Once confirmed, the environment is Copilot-ready.
8. **Proceed with Copilot adoption.** At this point, the organization has met the Phase One requirement in Microsoft's **Copilot for Microsoft 365 Adoption Playbook** of reviewing security and data settings. The organization is now in a position where it can proceed with the remainder of Phase One and subsequent phases of the Copilot adoption.

This workflow can be visualized as follows:



The four key outcomes of this workflow are in the middle of the diagram:

- a. Privacy data minimization,
- b. Privacy data consolidation,
- c. Privacy data labeling and
- d. Identifying privacy data not protected by Purview.

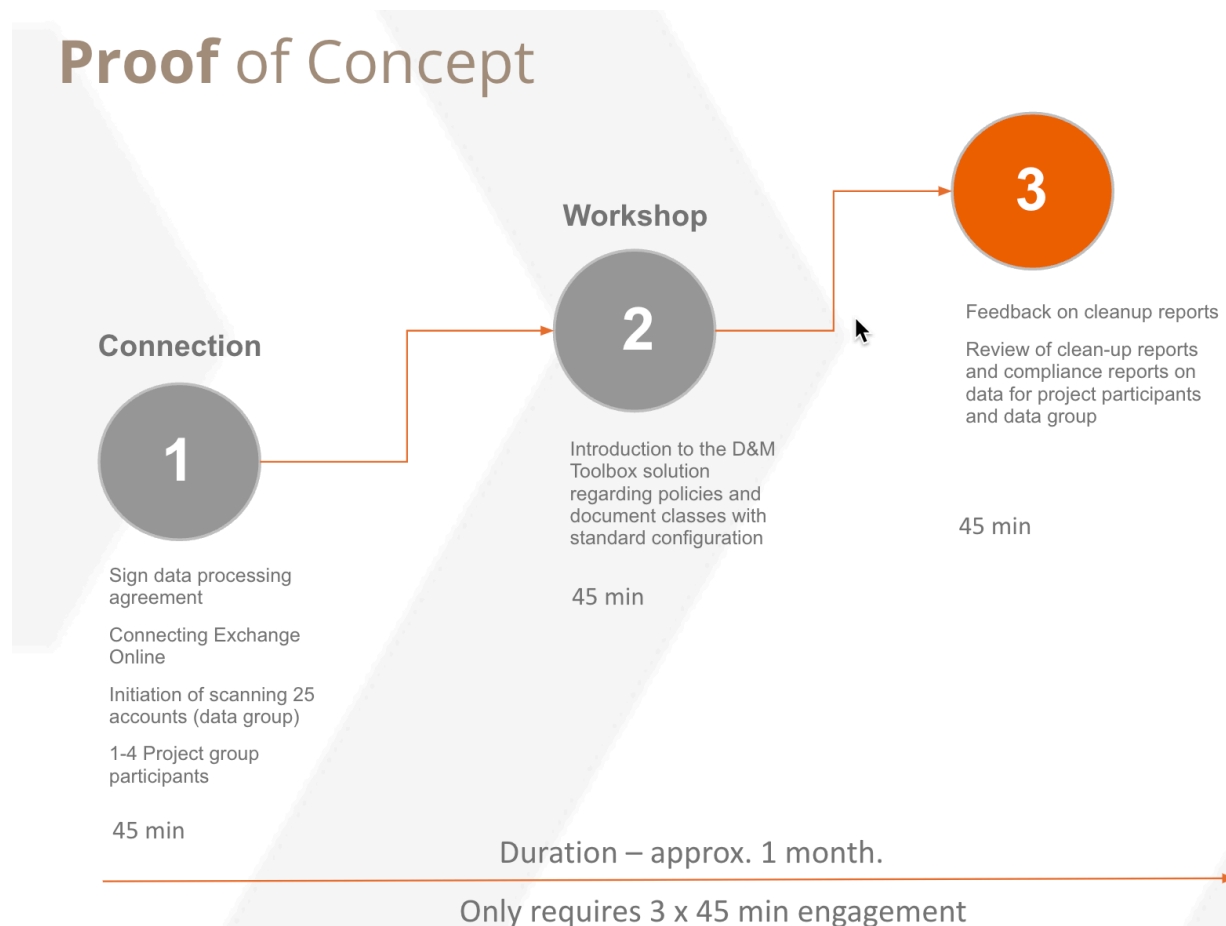
These items are facilitated, at scale, by **Data & More Copilot Privacy Protection**.



11 Get a Sponsored PoC

There is nothing like a hands-on evaluation of a white paper. If you want a demo of our software, you can try demo.dataandmore.com. It will give you five users out of the box.

If you send an email stating that you would like to evaluate Data & More, we will provide you with 20 more users to test and a free compliance workshop if your organization has more than 200 employees.



w: dataandmore.com // m: support@dataandmore.com //

Denmark: +45 4290 1070 - Flaeskertorvet 68, 1711 Copenhagen V, Denmark

Germany: +49 151 59422362 - Am Steinebrück 29, 40589 Düsseldorf, Germany

Canada: +1.587.966.9070 - 500 - 4th Avenue SW, Calgary, Alberta, Canada

