Data Risk Assessment by Forcepoint Management of Personal Data

Trial Version of Forcepoint DSPM and Forcepoint DDR

Forcepoint

MPD Guide

Forcepoint
April 19, 2025

Data Flow Summary

The Forcepoint classification pipeline uses an AI Mesh composed of multiple small models to analyze whether the data subjected to the analysis is sensitive or confidential. The mesh may contain keyword filters, AI sentiment analyzers , and mappings from intermediary classification outcomes to desired outcomes (ex. is this file Confidential).

The AI sentiment analyzers classify files by using diverse machine learning algorithms to analyze text, identify keywords, phrases and contextual cues from files. It does not use any biometric data.

The Data Catalog section provided in this document elaborates on the high-level summary, providing more detailed depiction of the underlying components and their various interactions.

A Data Risk Assessment (DRA) report is available as soon as a scan/a file is touched(DDR) will DRAstart which will identify the various kind of files(PII,PCI, risk, confidential etc) available in the Data Sources

Data Protection

Transfers
Only metadata will be shared from the Data source to DSPM.Policies can be set up to receive notifications and admins can manually take necessary actions on the files (e.g: Revoking permission for a user). Users registering for the trial program will have the connection setup with their Data source (One Drive) from the DSPM platform. This will be secured as we will be using an encrypted HTTPS connection for this purpose. No files or contents of the files will be shared from One Drive to DSPM.

At Rest/Stored
Forcepoint can attain swift and compressive visibility and classification for all data whether in motion or at rest.

Credentials
Customer SaaS credentials are encrypted using per-tenant asymmetric key. Customer onboarding credentials are shared via 1Password. On-Prem credentials are encrypted and encoded in the form of base64.

Data Collection Request

At certain times mutually agreed with the customer, we need to collect information regarding the performance of the AI classification software, with the purpose of tailoring the behaviour of the AI mesh to the customer needs. Typically, the information collected on these occasions is anonymized, then it is analysed by the Data Science team, and then results are shared with the customer as a report containing the suggestions on how to improve the performance of the AI. (We will not access the content collected without the customer consent.)

What we WILL collect: metadata associated with classification, including file names, date times created, classification outcomes from our own models, associated with the files classified by the classification product.

What we WILL NOT collect: any of the original file contents.

Anonymization: all the information except the metadata generated by the classification process (AI classification outcomes and their associated confidences) will be destroyed or anonymized on our side before analysis by the Data Science team. This will include, but will not be limited to, file names and paths (replaced with anonymous IDs), usernames, and any data signatures and hashes.

Retention and Deletion

The details (Name, email id) collected during the registration process of the DSPM trial will be temporarily maintained until end of the Trial program.

The tenant that is spun up and any associated data that is part of the trial will be deleted immediately as soon as the trial period ends(15days-30days).

Any account that is created in salesforce as part of the trial will be deleted after the trial period ends, if they do not proceed with purchasing the DSPM/DDR product.

Any GDPR Subject Access Requests (SARs) for deletion of personal data from HubSpot or Salesforce can be sent to privacy@forcepoint.com. After a file is evaluated, the per-file outcomes(metadata) from the classification network within the AI mesh are stored in a database, making them accessible to GQL enabled filters and reports.

Any metadata that is used for the Data Risk Assessment (DRA) report will be deleted at the end of the trial.

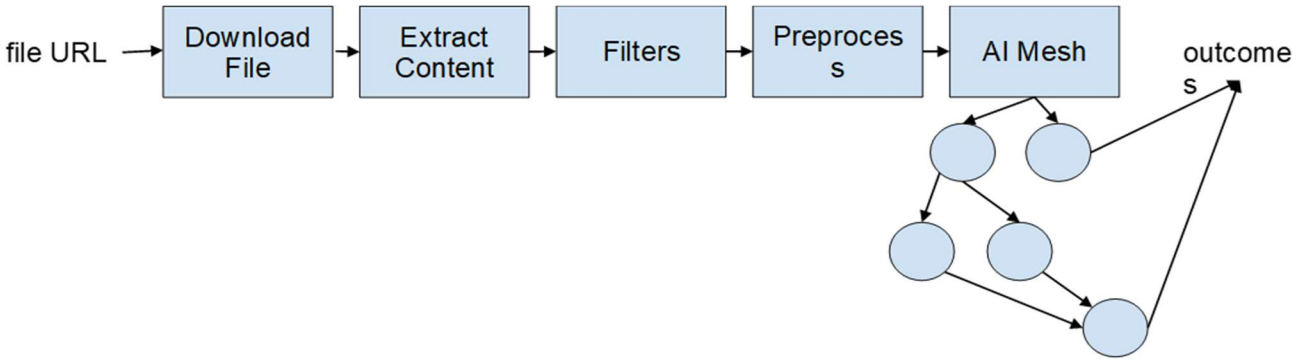
Prevention of Data Abuse

Admin access and logging of actions – Actions of who access the data can be tracked through Data lineage.

The classification pipeline incorporates Active Directory information about who has access to the files. This integration is important for assessing the risk associated with highly confidential files being accessed by trustees.

Data Catalog

Data Flow table(s) below provide a visual representation of where a data flow starts and ends, and the key product components involved.



The classification pipeline uses an AI Mesh composed of multiple small models to analyse whether the data subjected to the analysis is sensitive or confidential. The mesh may be organized as below, and may contain keyword filters, AI sentiment analyses mappings from intermediary classification outcomes to desired outcomes (ex. is this file Confidential). The AI Mesh contains narrow-AI / Machine Learning models which have been trained exclusively on text data that is available in publicly available datasets that are verified by internal data scientists and does not use large-scale, automated scraping of public data. The AI Mesh reads text extracted from the original files and processes and forwards it depends on which mesh nodes are available.

DSPM/DDR Incidents

Data Source	Data Content	Personal Data	Purpose
Microsoft OneDrive	Any data that the customer adds to the data sources (One Drive) and is scanned.	DSPM can detect various data types, including PII, PCI. By default, DSPM will scan all the data within the customers designated location that is provided. The Username and Email is stored from the web form to set up the account to One Drive	The incidents are available for an administrator to investigate for an intentional or accidental data leakage.

Data Processing Details	
Processing	*Incidents are created by DSPM upon discovering of files. The information of the incident includes file paths, classification, compliance tags, etc., which is displayed on the dashboard. Various policies can also be defined for alerting to be done on the files detected.* *A Data Risk Assessment Report (DRA) can be viewed or downloaded at any time once a scans starts or a file is touched DDR) to view various details like PII, PCI, risk, confidential etc* *The contents of the files in Data sources are never sent to the DSPM server, we only collect and transfer the metadata from the files that is obtained after inspecting it for details like confidentiality, classification, Compliance etc.*
Retention	*We do not retain any customer data files, and all metadata collected from the data source will be deleted when the trial ends.* *No files are transferred to DSPM. Files will be scanned within the Data source and will only return the metadata to DSPM,* *We will store information such as the name and email address in HubSpot and Salesforce for the duration of the trial period. If the customer decides not to proceed after the trial, this data will be deleted.*
Protection	*There will be a secured encrypted HTTPS TLS connection for the purpose of connecting from DSPM to the Data sources. No files or contents of the files will be shared from the Data sources to DSPM*
Deletion	*Incidents cannot be deleted from the application. However, an administrator can move the incidents to a quarantine location in the customers environment which can then be reviewed and deleted.* *At the end of the trial (30 days to 60days) all telemetry data and logs will be deleted including the account details username, email address.*
Pseudonymization	*No Pseudonymization is applied to the data. The contents of the files in Data sources are never transferred to the DSPM server, we only collect and transfer the metadata from the files*
Management of other Privacy Requirements	
SAR - Right to Access	*Customer has all access to the metadata that is collected by DSPM.*
SAR - Correction/Rectification	*The UI enables manual remediation of incident by providing options of sending files to a quarantine location. The administration can decide what is to be done with those files from that location.*
SAR - Right to be Forgotten	*The UI enables manual remediation of incident by providing options of sending files to a quarantine location. The administration can decide what is to be done with those files from that location. The administrator have the ability to delete the data locally after the files have been moved to the quarantine location.* *Once the trial period ends, all of the data is deleted.*
Data Storage / Localization	*After a file is evaluated, the per-file outcomes from the classification network within the AI mesh are stored in a database, making them accessible to GQL enabled filters and reports. All DSPM tenants are located in US East (Virginia).*
Cookies	*Strictly necessary cookies - Always active. These cookies are necessary for the browser to function with DSPM.* *Performance cookies: These cookies are used to improve the performance of DSPM and our website.,* *Functional cookies: These cookies enable the website and browser to provide enhanced functionality and personalization when using DSPM.*

Forcepoint

forcepoint.com/dra

About Forcepoint

Forcepoint simplifies security for global businesses and governments. Forcepoint's all-in-one, truly cloud-native platform makes it easy to adopt Zero Trust and prevent the theft or loss of sensitive data and intellectual property no matter where people are working. Based in Austin, Texas, Forcepoint creates safe, trusted environments for customers and their employees in more than 150 countries. Engage with Forcepoint on www.forcepoint.com, Twitter and LinkedIn.

DRA Trial Version of Forcepoint DSPM and DDR Management of Personal Data 22 Apr. 25