Forcepoint Data Classification

powered by Getvisibility

Combining generative AI and Large Language Models to deliver highly accurate data classification for Forcepoint Enterprise DLP and Forcepoint ONE

Fundamental Challenge to Data Security

When it comes to an organization preventing the loss of data, the common phrase, "you can't protect what you can't see," speaks to the very foundation of data security. Organizations that lack an effective method for identifying or "seeing" all the types of data that they have are ultimately unable to prevent the loss of that data. However, typical classification methods are highly manual, largely depending solely on individuals to make important decisions on the value of data. Inaccurate data classification generates large volumes of false positives and false negatives in Data Loss Prevention (DLP) policy enforcement. This results in wasted time and resources. It also impacts an organization's ability to keep sensitive data from being exfiltrated and leaves it open to security threats from both outside (malware, ransomware) and within the organization (malicious insiders, misuse, error, mistaken classification).
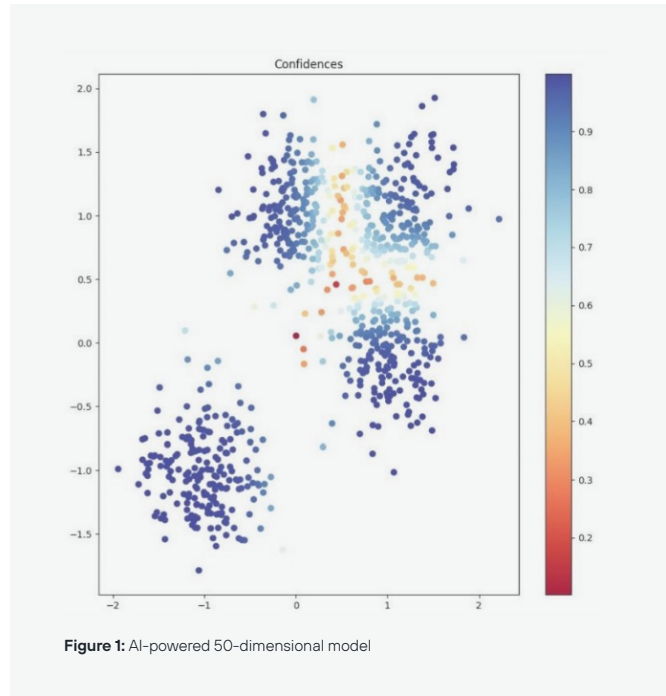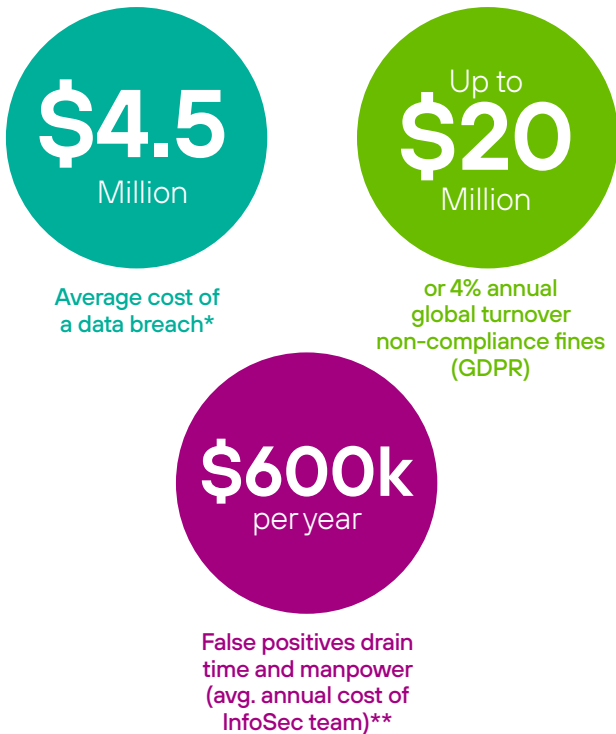

Figure 1: AI-powered 50-dimensional model

Impact of Inaccurate Data Classification

$4.5 Million

Average cost of a data breach*

Up to $20 Million

or 4% annual global turnover non-compliance fines (GDPR)

$600k per year

False positives drain time and manpower (avg. annual cost of InfoSec team)**

* Source: IBM Security: Cost of a Data Breach Report 2023
** The Quest for a Single Set of Unified DLP Policies (IDC)

## Leveraging AI to Gain Sight into All of Your Data

Forcepoint is delivering an innovative type of data classification technique that leverages the power of Artificial Intelligence (AI) and Large Language Models (LLMs) to improve data classification accuracy, automating the discovery and classification process.

Moving beyond traditional machine learning, Forcepoint Data Classification employs TensorFlow to create Generative AI models, constructing a predictive and self-learning classification engine. Forcepoint Data Classification works with natural language to suggest classifications. Using a 50-dimensional model, customer meta-data is ingested into the model and, through continuous learning, delivers high-accuracy suggestions.

To start, thousands of newly created files and emails are classified for customers. Each file is converted into a mathematical vector with a classification label that is anonymous and not able to be reverse engineered so that the information can be securely stored. In this way, Forcepoint Data Classification leverages classification from customers across various industries to build and continuously improve the 50-dimensional model. As an individual organization deploys and runs Forcepoint Data Classification, they are also able to maintain their own 50-dimensional model that can learn based on their unique organizational classification requirements.

Forcepoint Data Classification achieves unparalleled accuracy by combining synthetic and real-world data:

› **Leading LLMs generate millions of diverse sample files,** ensuring a comprehensive training dataset that spans a vast range of content types, industries, and scenarios.

› **This synthetic data seamlessly complements anonymous customer data,** training the 50-dimensional model to learn and classify with remarkable accuracy.
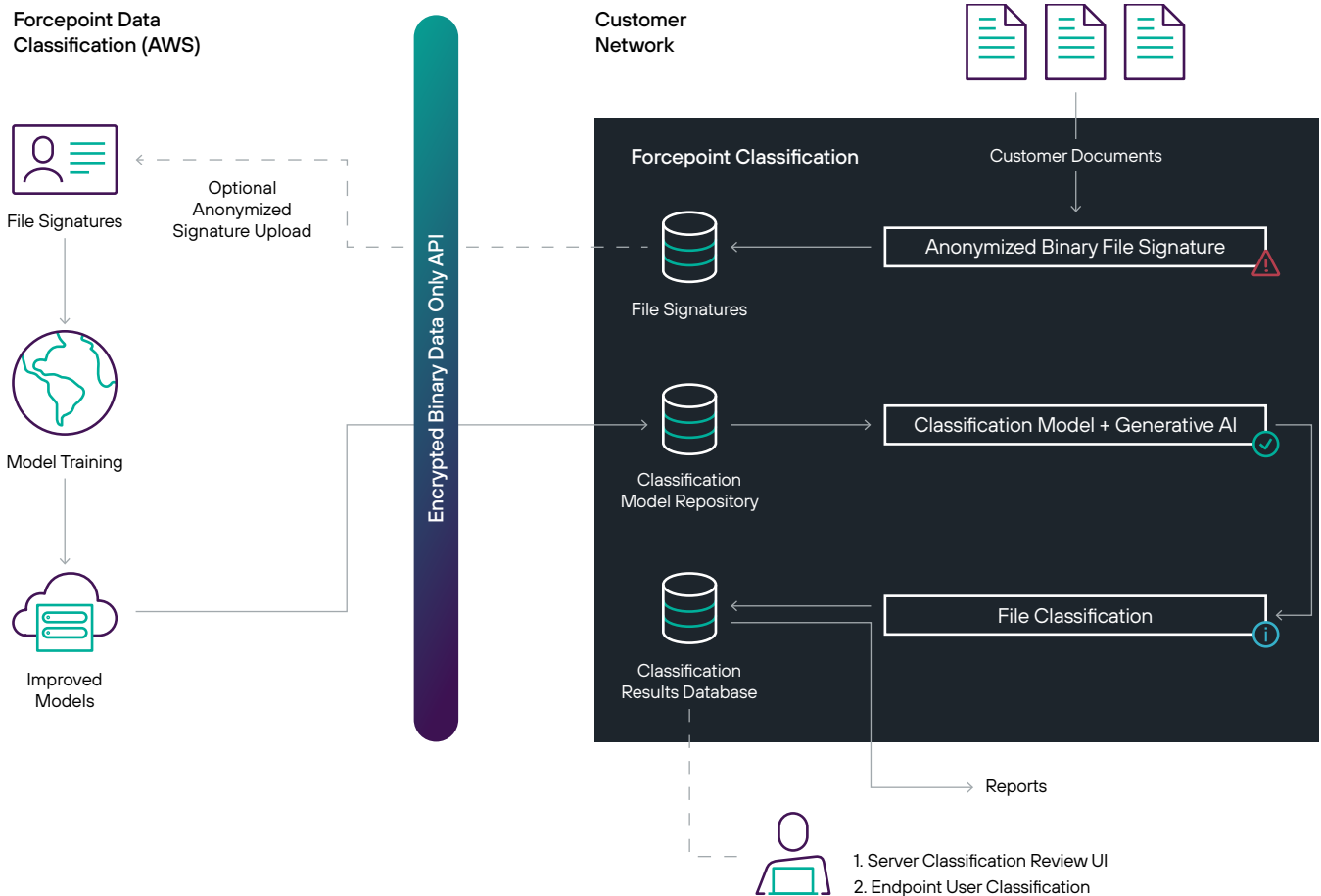
During file creation, Forcepoint Data Classification enables you to add visual labels and metadata, giving you the option of selecting your own classification or using the AI-generated recommendation.

Each recommendation is delivered with a confidence level based on the data within the file compared with the AI model to help make highly accurate classification decisions. The decision then enables you to add visual labels and metadata to facilitate policy enforcement and regulation compliance. As the tool is used, it continues to learn and improve. The customization configuration wizard options include:

› Compliance
› Classification
› Enforcement rules
› Visual labeling
› Email sharing rules
› Exceptions

The outcome is Forcepoint Data Classification delivers "Classification Everywhere," providing on-the-fly classification of newly created data anywhere your people work. Through continual learning and improvement, it delivers the most accurate classification tailored to your business needs, all while requiring minimal effort to deploy and ensuring a superior Total Cost of Ownership (TCO).

› **AI-powered** for industry-leading precision and efficiency

› **Seamlessly deploy** with no user training

› **Increased productivity** and reduced false positives

Seamless Industry-Leading Precision and Efficiency
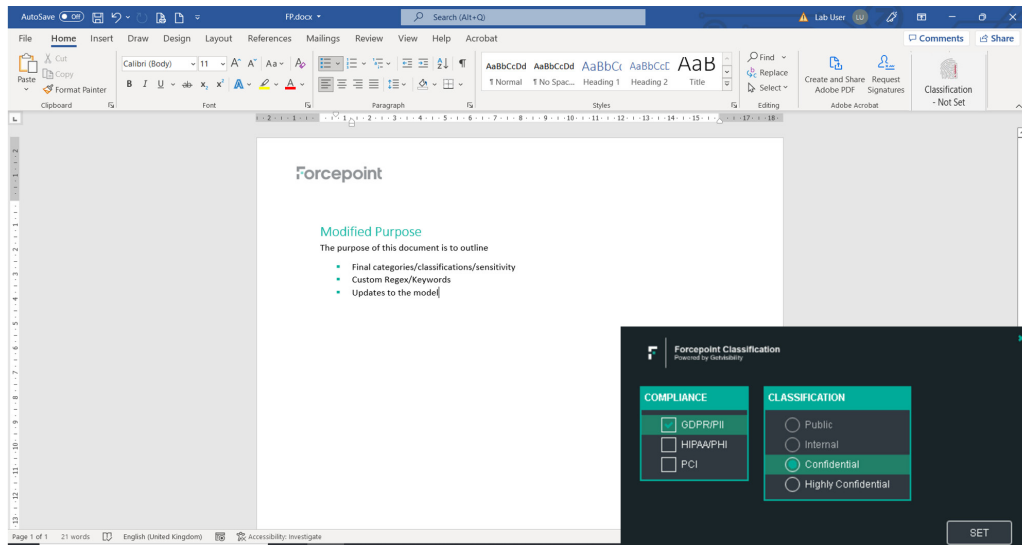
Forcepoint Data Classification is an industry-leading solution that seamlessly sits in the Microsoft application ribbon, providing easy access to the current document classification status. Upon selection, users receive prompts to choose classifications with AI-powered suggestions. After saving, the chosen classification level is automatically applied and embedded in the document's metadata.

Accuracy

Forcepoint Data Classification uses advanced AI to suggest classifications to users, dramatically increasing accuracy and reducing human error.

Automation

Automate the classification and protection of your data in real-time as your teams work and collaborate on apps anywhere.

Visual Labeling

Customizable visual labeling allows you to add custom headers, footers or extra text fields, based on your organization's policies.

User Activity Reporting

The administrator dashboard lets you track user classification activity, pinpointing areas for policy updates, employee training or potential insider threat investigation.

Compatibility

Seamlessly integrates with a wide array of applications, including the Microsoft Office suite and more, and pairs with Forcepoint Enterprise DLP and Forcepoint ONE, enabling immediate AI-powered data classification upon deployment.

Streamline Compliance

Forcepoint Data Classification makes it easy for organizations to meet compliance standards. The platform simplifies the compliance process with a user-friendly customization wizard, ensuring adherence to regulatory requirements. Built-in selections for major regulations further simplify compliance audits and reporting. Continuous data monitoring ensures ongoing adherence with evolving regulations so your organization can focus on innovation and business outcomes.

› Out-of-the-box key regulations simplifying compliance

› Continuous data monitoring ensures ongoing adherence to evolving regulations

› Granular policy controls allow you to tailor-fit to specific compliance needs

› Streamlines compliance with the broadest coverage of data types in the industry

For more information schedule a demo today.

Appendix A - Key Features:

FEATURE	OVERVIEW
Data Classification	Windows and UNIX File Servers, Windows Unix and MacOS computers, Android and iOS mobile devices, NAS, Microsoft Online, DropBox, Box and GDrive (uses AIbased classifier).
Personal Identifiable Information (PII) Identification	AI named entity recognition models which identify PII based on the text content, offering the best PII identification accuracy available on the market.
Email Control	AI for email content and attachments classification. Allows keeping track of communication that involves sensitive subjects.
Monitoring	Monitoring and recording of events related to file activities and changes to active directory AI anomaly detection.
Reporting	Widgets provide a high-level overview with the ability to drill down and conduct detailed forensic investigations. All the findings can be exported in various formats. Generates management executive reports with high-level overview of data risk posture.
File Type Support	Supports more than 50 file types (including Microsoft Office file types) including the following standard files types: pdf, doc, dot, xls, xlt, ppt, pps, docx, docm, dotm, dotx, xlsx, xlsm, xlst, xlsb, pptx, potm, potx, ppsm, pptm, ppsx, vsdm, vsdx, vstx, vss, vssm, vst, vstm, vssx, dwg, dxf, jpg, jpeg, png, mp4, jpe, bmp, wav, mov, avi, m4p, m4v, gif, tiff, tif, zip, mpp.
Internet Dependency	Staff can classify and tag documents using the same rules as when they are online with the same warnings, blocking of risky activities and help assistance to explain the reasoning for the restrictions. While the ML classification suggestions are only available when online, the pattern-based suggestions are continuously available offline as well as online.
Integration Capabilities	Out-of-the-box integration with Forcepoint Enterprise Data Loss Prevention (DLP) and Forcepoint ONE Integrated DLP. It also is able to be used with other leading DLP solutions in the industry.
Rights Management Partner Integration	Out-of-the-box integration is available with the main RMS providers. This includes Microsoft Azure RMS, Seclore, SealPath and Ionic.
Visual Marking Customization	Visual marking and metadata is fully customizable supporting different attributes and variables.
Existing Header/Footer Management	Visual markings are flexible in terms of position. Values can be inserted in existing headers/footers or can replace old headers/footers if needed. Watermarks are also supported.
Comprehensive Reporting	Reporting does not require additional resources to be enabled on the client machine that could lead to reduction in performance. Logging is stored on a centralized server. Forcepoint Classification will provide reports on user activity, data at risk, risk scoring and risk by department.
Email Subject Marking	Fully supports subject marking on emails.
Content Checking on Embedded and Attachment Files	Fully supports content checking on embedded as well as attachment files including ZIP files.

Forcepoint

forcepoint.com/contact

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.