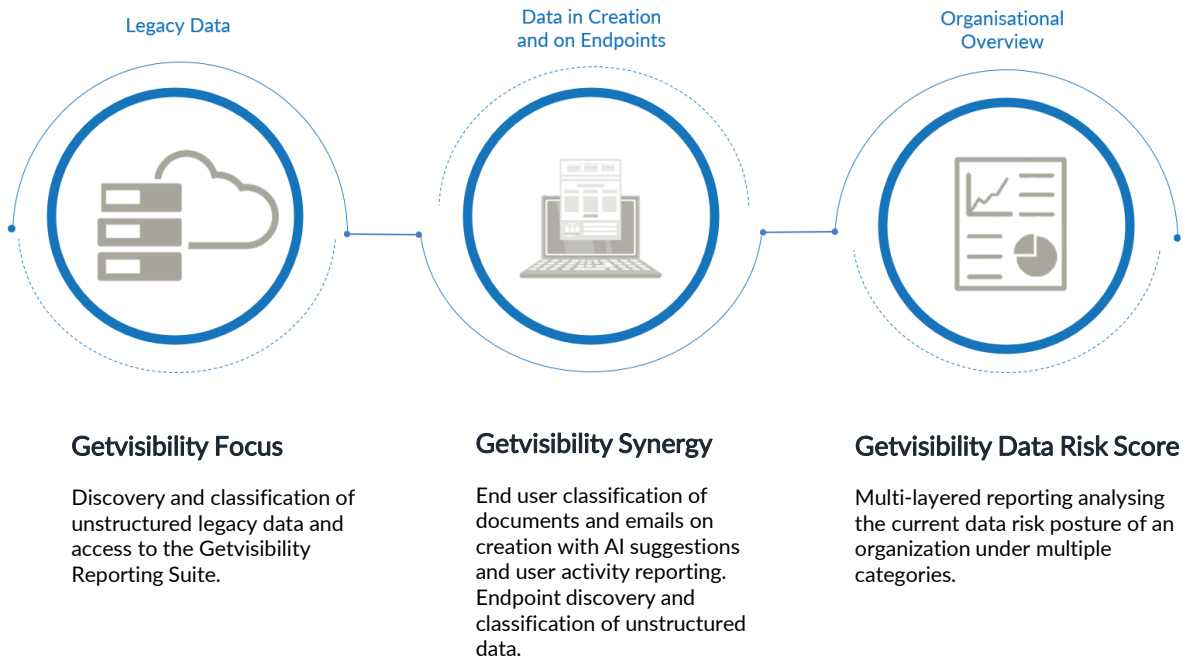# Data Discovery and Classification

Organizations have enormous quantities of data scattered across file servers, cloud infrastructure and mobile devices. Almost every organization will have files on their network that contain sensitive, personal or regulated content. Unfortunately, most organizations have no idea what their unstructured data footprint looks like and they're creating significantly more data on a daily basis. Getvisibility's approach is to help companies understand their data footprint with the highest accuracy possible and control this data. Without this, unstructured data security simply cannot work.

Getvisibility has created a solution using the latest in artificial intelligence to help companies to protect and govern their data like never before, in a quick and cost effective manner.
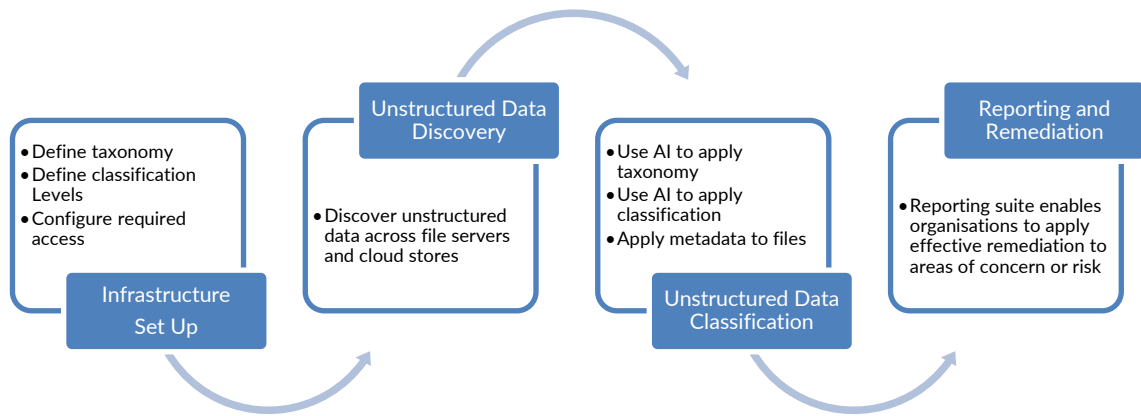
Getvisibility's dual approach to data classification, allows end users to classify documents as they are created, aided by AI suggestions. Once classified, the suitable level of protection is then applied. Our AI engine then verifies end user activity and provides reporting on misclassification and unusual activity. Meanwhile, using state-of-the-art machine learning algorithms, we combine natural language processing with neural networks and optical character recognition. This allows us to discover and classify unstructured data across organizations with unparalleled accuracy and speed without affecting server performance. The Getvisibility reporting suite enables data remediation through tactical reporting on issues such as duplicate data, redundant obsolete or trivial data (ROT), data retention, access permissions and much more.

Key Elements

Legacy Data	Data in Creation and on Endpoints	Organisational Overview

Getvisibility Focus

Discovery and classification of unstructured legacy data and access to the Getvisibility Reporting Suite.

Getvisibility Synergy

End user classification of documents and emails on creation with AI suggestions and user activity reporting. Endpoint discovery and classification of unstructured data.

Getvisibility Data Risk Score

Multi-layered reporting analysing the current data risk posture of an organization under multiple categories.

Steps to Effective Classification

Effectively classifying unstructured data requires a number of simple but crucial steps. This process needs to begin with the organisation clearly defining both their taxonomies and data classification policies. These are customisable as per the organisations specific needs. Once this step is completed and the technical configuration is completed, data discovery can then begin. The data discovery process allows the Getvisibility server within the organisation to scan unstructured data across multiple file stores both on premises and in the cloud, providing a comprehensive inventory of this data. Getvisibility's advanced machine learning models then assign these file a category, subcategory and classification level in line with the companies pre defined policies. For example, HR Document, Appraisal, Confidential. Once documents have been classified, users can generate reports via the reporting suite or view detailed results through the *Results* tab on the Getvisibility dashboard.

Sample Results