

GDPR Content Identification Service

Understand and Classify your Data.

PII Discovery and Identification.

Determining the amount and location of sensitive information across multiple unstructured data repositories is a significant undertaking and typically serves as a barrier to accomplishing this task. The Content Identification Service (CIS) allows organisations to take the first steps into discovering and understanding the amount of sensitive Personal Identifiable Information (PII) that exists within their unstructured data repositories. CIS provides an invaluable insight into the size and scale of your GDPR data compliance challenges, identifying sensitive information that is central and relevant to the success for GDPR compliance in a streamlined and automated fashion.

The CIS provides a detailed analysis and report, based upon content of the ingested data from an identified data sample (Maximum 8TB). The findings from this analysis will provide an invaluable insight into the level of data file duplication, data types, individual file size and aging characteristics, whilst also classifying data items that contain PII/sensitive information.

Service Deliverables.

CIS comprises of two key phases which are outlined below.

Phase 1: Understand and classify your data.

Discover what volume of PII resides and is located within the ingested unstructured data repository under study.

Phase 2: Take action.

Once the location of the PII information has been identified, additional analysis will be performed to further categorise the data and highlight the potential GDPR compliance challenges

Action 1. Reduce your data management challenges (and associated storage costs and liabilities), by identifying Redundant, Obsolete, and Trivial (ROT) data items.

Action 2. Sort & Categorize the remaining data into information that is relevant to GDPR and therefore “in scope”.

Action 3. Further categorise this data into simple risk domains i.e. low, medium and high

Service Activities.

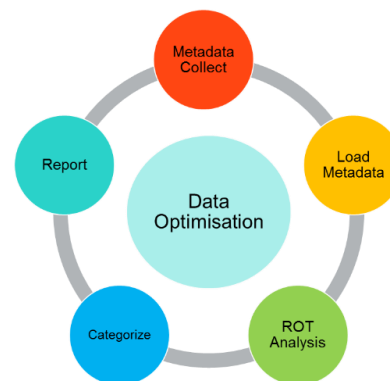


Figure 1. Data Optimisation Methodology.

Delivery of the CIS service requires the following tasks/actions and deliverables:

1. Pre-Installation meeting/ conference call. This is generally a technical exchange/data gathering conference call to agree the technical requirements.
2. A Statement of Work (SoW) will be prepared which describes the project methodology, approach and software implementation and data collection process.
3. The SoW will include references to the Microsoft Windows data repository locations where the source unstructured data resides which will be ingested for analysis.
4. The SoW will outline the technical requirements for the CIS implementation i.e. VMware clients, local storage, Operating system versions.
5. Post the on-site software installation and configuration. The data ingested process will be initiated.
6. When the data ingestion and content indexing has successfully completed, StoreXltd will undertake.
 - 6.1. High-level data analysis (i.e. ROT, data type, growth...).

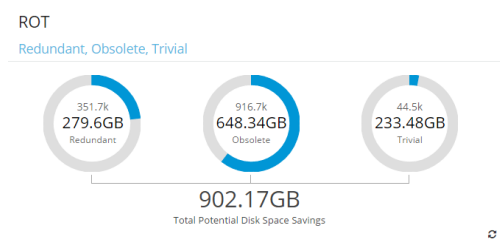


Figure 2. Data ROT Analysis.

- 6.2. PII and Sensitive information categorisation.

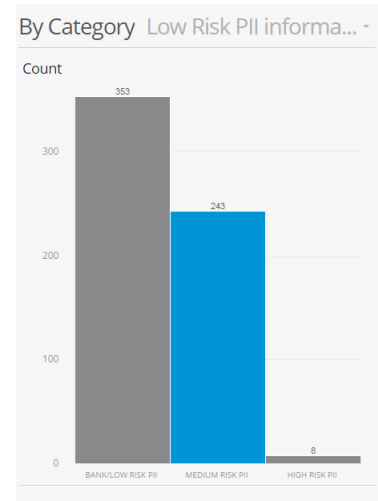


Figure 3 Low, Medium and High Risk PII Classification.

- 6.3. Identified data items should be considered “In-scope” as part of the GDPR.
- 6.4. Determine level of “structured” vs “unstructured” data.
- 6.5. Determine level of file duplication.
- 6.6. Identify amount/level of obsolete aged data files.
7. Findings and recommendations to be presented and discussed with the customer after the written report has been received by the customer.

Service Exclusions.

It is important to note that this service specifically excludes the following:

- Any remedial activity. However, if assistance is required to address some or all of the findings, it is suggested you request details of additional packages or consultancy services

- Any configuration changes.
- Any Capacity Planning or detailed monitoring of the system.
- A hardware/patch audit of the system.

Delivery Pre-Requisites.

To ensure this service is delivered efficiently and effectively it is critical that the customer.

- Ensures that technical assistance is made available as and when requested by the technical specialist performing this study.
- The customer identifies the single (One) data repository (UNC path) that will be indexed.

Customer's environmental Pre-requisites.

As above, also including:

- Availability of a Windows Service Account which can be associated with the associated installed software services.

Software Pre-Requisites.

- 3x Vmware Virtual client machines each with Windows 2008 R2 or Windows 2012 installed and configured.

Hardware Pre-Requisites

- Sufficient local storage to host the ingested data (typically, 5% of the volume of the Ingested data)