DataCube

Email: info@apperception.co.uk
Tel: 0845 644 3479
Web: www.apperception.co.uk/datacube

Data Cleansing and Data Retention

- Forensic Information Auditing
- ROT (Redundant, Obsolete or Trivial) Identification & Removal
- Understand what is held and where
- Identify target content, joint action plan and remediation lists for data migration, quarantine or deletion
- Aged File Analysis
- Information Remediation
- Culling, reorganising, archiving and migrating as appropriate

Organisational Requirements

Understanding what data you are actually holding, for better or for worse, is critical. Data management is an increasing problem that will only become progressively more difficult to address the longer it is left. Organisations need to be able to understand exactly what they have in their network storage. Much of the data stored will be redundant, trivial, duplicatory, orphaned and even illegal in some cases. The ability to accurately identify what information can safely be disposed of, quarantined or transferred to lower cost data storage has, until now, been beyond almost all organisations.

DataCube Solution

DataCube uses ground-breaking conceptual search technology to enable customers to see what they really have - to identify duplications, ROT, cleanse and categorise their data to enable either controlled reduction of volumes, migration to other systems or combined solutions.

DataCube creates a Data Inventory of every documents location, its metadata details and its content categories. This enables it to be categorised against the organisations taxonomy or schema, thus enabling the inclusion of Data Retention dates to be tagged and the protective marking levels of that document within that category to be recorded. Action can then be taken to fulfil the Policy requirements. At no time does the document itself move location nor does the existing metadata (i.e. last accessed) change.

DataCube

Email: info@apperception.co.uk
Tel: 0845 644 3479
Web: www.apperception.co.uk/datacube

Features

- Conceptual search capability analyses millions of documents at lightning speed - much more accurate and effective than keyword searches
- Identifies all file types and their network location
- Redundant, Obsolete, Trivial data quarantined for removal
- Duplicates and near duplicates quarantined for removal
- Identification of Illegal Files (Music. Images, YouTube clips, private software)
- Orphaned data - generated by people no longer in the organisation or old projects
- Inaccessible data- corrupt, old permissions, altered rights
- Low/zero usage data or application recall

How Does it Work?

The DataCube connects to Active Directory and discovers the devices, users, folders and files contained within the organisation's (target) network. A Data Collection Management application (which is configurable) is then run to scan all of the required file types for the specified domains in order to build a Data Inventory. As part of the data collection and build of the Data Inventory, DataCube can perform a MD5# calculation on all of the data to identify duplicate and near-duplicate files. Each Data Inventory record records whether the file has a duplicate and provides a list of all the duplicate file locations.

The Legacy Data Management Reporting option within the DataCube Dashboard then provides a range of reports which can be used in relation to ROT data. This includes duplicate files, illegal files, corrupt files, and text files with little or no content and .tmp files, as well as reports using the metadata such as Last Accessed Date (showing redundancy) or even matching orphaned user files. As the DataCube can analyse and understand the content of the data, the system is also able to add business rules into the searches (for example trivial content such as meeting agendas, car park arrangements etc.)

Organisational Benefits

- Identifies data ready for quarantine, disposal or migration to a lower cost storage regime. It removes the duplicates, ROT and cleanses legacy data, freeing up space and enabling other activities like Retention Policy enforcement and Retro labelling to be undertaken.
- Reduces cost of storage and operations by decluttering - by up to 45%
- Enables an understanding of the data held to be gained. Discovers sensitive data's whereabouts and either disposing of or removing to a safer environment.