Data Mapping - Revised Definition

Data Mapping is the process of identifying an organization's data sources; and understanding:

- how those data sources are stored, structured, managed and accessed;
- how those data sources and the data are used within the organization;
- who is responsible for managing those data sources; and
- the applicable retention and back-up practices and policies for the data sources.

Data Mapping can be carried out in response to litigation, where data relevant to a specific matter is sought. However maintaining an accurate Data Map across your organisation is important for information governance, records keeping, data privacy, among other issues.

The Data Mapping information is often detailed or explained in a Data Map document or visualisation. Once created, a Data Map needs to be updated whenever a change is made to any of the elements that constitute the Data Map.

There are different definitions of Data Mapping depending on the context. From an IT perspective, Data Mapping is the process of locating files, and suggesting associations between files or user generated content within a large body of data - which may or may not be apparent using other techniques. For example, a team may create a Data Map of all the locations where they save data, what type of data is saved there, and which other teams have access to it.

Why is Data Mapping important?

As data volumes grow, along with the ongoing introduction of new technologies and information repositories and increasing legal and regulatory requirements, it has become more important for companies to understand what data they have and where it resides to aid in the seamless protection or extraction of data.

To leverage data and extract business value out of it, the information collected from various external and internal sources must be unified and transformed into a format suitable for operational and analytical processes. This is accomplished through data mapping, which is an integral step in various data management processes.

Through the process of data mapping, knowledge is garnered about what is stored where, so that compliant, defensible disposition programs can be established and executed, including the implementation of legal holds and any subsequent eDiscovery requirements that may arise as a result.

Data mapping is about maintaining relationships between data elements and helps in identifying and securing private and personal data. For example, data mapping helps in securing information like bank details, social security numbers, health information, etc. It also aids in identifying and resolving data exceptions. Data mapping can also establish standards, such as patterns of naming conventions (e.g. a date nomenclature should incorporate "mm-dd-yyyy").

In addition, good data mapping practices lay the foundation for building information-related projects and programs and also better supports system/data migration and decommissioning initiatives. In essence, it allows you to effectively use the data map to bring consistency to data and leverage an organization's information assets.

Data mapping is among the most important design steps in data migration, data integration and business intelligence projects. Mapping source to target data greatly influences project success – perhaps more than any other task. The outcome of the mapping process is a primary tool for communications between project architects, developers, and testers.

A data map is also a very useful tool in the event an organization is faced with searching for responsive documents as part of an investigation or litigation as it allows an organization to easily locate what is stored where on an employee by employee (or custodian by custodian) basis. Data mapping allows organizations to store and transfer large amounts of data in a quick period of time with the added benefit of analyzing the root causes of any abnormalities. It also lowers the cost of business operations and assists in customizing any collaboration effortlessly and securely for all team members.