How do you create a Data Map?

Core Elements to Creating a Data Map:

1. Consider the use case for the data map. For instance, are you creating a data map as an information governance tool on behalf of an organization or instead in response to an investigation or litigation or alternatively to ensure regulatory privacy and compliance? The use case will assist in guiding the content that should be captured within the data map as well as the stakeholders that should be involved in the effort.

2. Identify the stakeholders for data map creation. Depending on the type of data map being created, you will potentially have to draw upon individuals across many different groups in order to assemble the requisite information.

- By way of an example, if you are creating a data map to track information about what data is stored where on behalf of an organization to ensure privacy and compliance, you will likely need to involve individuals from the following groups initially: IT, Legal, Information Governance and Third-Party Service Providers.
- Alternatively, if you are creating a data map in response to a corporate investigation, due to the confidential nature of the exercise the messaging surrounding the creation of the data map as well as the stakeholders involved would generally be limited to select individuals within Legal and IT.

3. Consider the disparate sources and location of data. There are both typical and atypical sources of data such as user-generated data or data that is system generated. In addition, there is content managed directly by an individual or an organization either onsite or in the cloud as well as third-party managed or hosted data. Depending on your organisation, data may be stored overseas or in other jurisdictions.

- As an example, if you are creating a data map in response to a litigation the scope is generally narrowed to a collection on behalf of pre-determined custodians.
- While conducting a custodian questionnaire will aid in assessing where potentially responsive documents might be stored, depending on the nature of the case, top data sources generally involve emails / email servers / email back-up repositories; personal and/or network shared drive data / network back-up repositories; documents stored on laptops or devices that are not otherwise synched to a network; smartphones; social media content; hard-copy documents etc. In light of the above, you will want to account for all of the data sources within the map.
- Alternatively, if you are creating a data map on behalf of an organization in an attempt to have a better handle on information governance, the scope and requirements of the map will be much more broad encompassing both typical and atypical data sources organizational wide such as mapping all of the applications used by different groups or potentially determining the location of printing routers.

4. Determine who owns the data

- There are both business and technical owners of data. On the business side, these are generally the individuals that create, share and/or distribute documents in the course of day-to-day business. On the technical side, these are the IT teams that typically interface with data on the back-end by managing items such as access, security, storage, preservation and retention and deletion of data. It's important that the data map incorporates ownership of information for both groups, as well as providing details about those that access the data and those that shouldn't have access to the data. Delineating the ownership information will aid in the successful preservation, retention and deletion of data as required.

5. Consider future updates (if required) to the map and determine who can be responsible for ensuring that the map is kept up to date. Instances such as personnel arriving or leaving an organization; decommissioning of certain data sources; implementation of new data sources or the integration of new regulations or laws can all potentially trigger an update to the map.

Avoiding Potential Pitfalls:

- *Plan Ahead:* Ensure that you are capturing all the requisite information in your data map the first time around so that you do not need to revisit gathering additional elements at a later date. Involving all the requisite stakeholders up front should assist with this task.
- *Time:* The process of creating a data map can be time intensive. Even the effort of obtaining buy-in from the right stakeholders will add to the timeline. If the creation of the data map is extremely time sensitive, then consider prioritizing ensuring the correct data retention and preservation measures are in place to avoid spoliation.
- *Third-Party Hosted/Managed Data:* It can sometimes prove challenging to wrap your arms around how data is stored, managed and access when third-party arrangements are in place. Most organizations will have some sort of integrated third-party hosted or managed applications or services however, so it is useful to think about this in advance as a discussion point with the stakeholders.
- *Local/International Applicable Regulatory Frameworks, Guidelines or Laws re Handling Data, including Confidential, Commercially Sensitive or Personally Identifiable Information (PII).* Consider what sort of information needs to be tracked in order to satisfy local or international · regulatory frameworks, guidelines or laws with regards to the handling of data so that this information can be tracked within the map, if required.

- *Ask First:* If you are generating a data map on behalf of an organization, you should inquire whether or not the IT group has a network map or network topology that you can expand upon for the sake of creating the data map.
- *Housekeeping for the Map:* Once created, ensure the map is a centrally located resource so that an organization can easily locate data. There can be implications of spoliation/costs if it takes too long to locate relevant data sources.