Discovery Data Collection Part 2

For the initial data collection discovery call, it's important to have representatives from the client who can speak to both the unstructured and source connection requirements to ensure there is a common understanding based on any feedback from the initial kickoff call and the data collection documents that were provided. We recommend building out a sample to help talk through and show these client requirements, recommend creating an overview on which data points originate from which kinds of sources like a source system or unstructured sources In a kind of template breakdown that articulates and highlights the makeup of each. When referring to unstructured data, we suggest leveraging this template to break out the topic, the data provider, potentially even the data approver as well and with flexibility of account of the data points to help dimensionalized the size of the content that was being captured. You can potentially pivot this information to rearrange it in a different set of columns, but this would be a good example where we've outlined what kind of data collection source if it's structured or unstructured and also the kinds of scenarios that we would be going through in order to refine the process and thus setting the stage in this discovery session.

If a claim has template in Excel, it is helpful to focus on how their existing template can be optimized via the data collection model and the experience if they do not yet have an existing template for example, potentially emails make up the majority of the current process, they would need to provide a list of metrics and metric names and the associated descriptions that can be used to define each captured data point. Showcasing a sample overlay sheet is also helpful when discussing these topics with the clients in order to highlight if they would be aggregating data potentially across different kinds of sites or different kinds of data input, ensure how these values would then be mapped, reported to a flat file sample that is ready to be consolidated into the data collection table. At this stage. It may not yet be needed as part of the initial discovery call unless it's highlighted as a priority during the kickoff call, but it is helpful and beneficial to start setting the stage for the customer to understand data organization and how this is important as we think about not only the discovery but the design and ultimately the build of the data collections within Workiva.

After doing an initial review of the documentation and files provided by the customer, we recommend assessing for any specific requirements to include in a sample template with an overlay sample highlighted in this field. Namely, is there historical data and comparisons across this unstructured template to help providers verify the values of providing or if it's just more of a historical data loading point, comments descriptions are additional notes for each data point. Also that help give context to the data providers or perhaps an entry feel that they are already accustomed and familiar with. We also recommend at a high level discussing current framework mapping approach and see if there are any additional frameworks beyond the core four plus CDP that are part of the base full solution. We recommend creating one template per scenario that is a priority based on the individual customer teams specific case. As we highlight a conversation around historical data, we recommend highlighting to the customer a sample of data and how the data would need to be formatted in order to load effectively into a Wdata collection table in order to populate historical information that would not be going through the data collection process. Here, we can highlight common instances that we've seen with some clients where there is metric information, potentially a section or a topic across that information with specific values that are showcased different columns for each year. In this example we recommend highlighting to the customer that in order to successfully load historical data information, the way to provide this content is to place it in a structure where the columns are fixed and unchanging and there are specific fields for both the year, the value, in this scenario, the units and all of these rows are stacked on top of each other to make it a structured consumable data format that can be loaded into the data collection table. We also recommend discussing with the customer if there's any source data connection that comes from structured external systems via that by asking for a sample CSV file loading this CSV file and creating a simple version of a Wdata table that matches the construct of the CSV file.

In the example here, ensuring the structured data source example, we've loaded the CSV file provided by a sample customer with six columns that are mapped from the content within that file that we've provided here very quick high level preview from that preview, we've set up a query that essentially just reads the table and by having created this query we can then showcase how this information would be consumed and flow through to the spreadsheet for a high level review but also ensuring that the columns that are provided are consistent with those that would be needed by the data model for a specific topic. What we're seeing here is a connection of the external incoming connection that actually showcases that CSV file if there was additional data that would be flowing in call it from that same source system could be updated via the connection, but this usually helps us the stage if there is potentially data transformation needs that need to be addressed, does the data need to be aggregated further in order to meet certain output requirements and also understand how this data would be consumed when staging later down the road in the factbook.