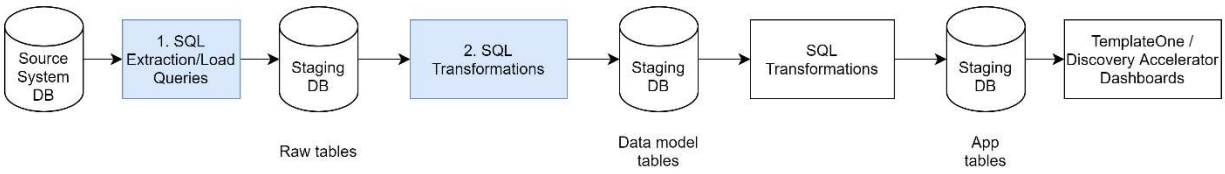
Appendix B – Data Connector Design Spec from Engineering

The output/deliverables consists of three parts:

1. Extract and Load: copying the raw data from the source system into a database.
2. Transform: transforming the raw data to the required data model.

3. BPMN reference models: for future usage, provide reference models for the process you are creating a connector for.



Extract and Load

The *extract and load* step copies the data into the staging database for transformation. In this step, the data is kept in the same form as in the original database. Transformations should not happen in this step but in a separate step.

Requirements:

- The queries for getting the tables should be in an appropriate form for the system:
    - For databases this should be SQL queries. It can be assumed that there is tooling for loading the SQL queries into the transformation database.
    - For SAP ECC6 this should be ABAP queries. It can be assumed that there is tooling for loading the ABAP queries into the transformation database.
    - For other systems the form needs to be agreed upon with UiPath. No tooling can be assumed for other systems. A description needs to be provided how the data is loaded.
- Columns that are not needed should not be loaded.
- The queries do not directly contain credentials or other secrets.
- The queries contain parameters for filtering on date ranges and concepts (like country codes)

Transform

The *transform* step transforms the copy of the source systems input tables to the required output (data model). Depending on the process, the data model either means the generic TemplateOne data model, or in the process specific models (currently P2P and O2C).

Link to TemplateOne Data Model

https://docs.uipath.com/process-mining/docs/input-tables-of-templateone

In this data model, the Tags and Due dates tables, and some separate attributes are not mandatory. For this RFP, however, they are part of the deliverables.

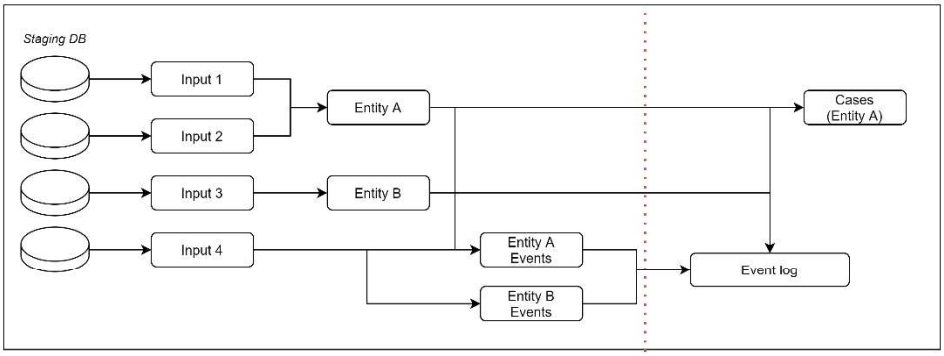
Link to P2P Data Model (for the P2P Discovery Accelerator. Process-specific)

https://docs.uipath.com/process-mining/docs/input-tables-of-the-purchase-to-pay-discovery-accelerator

Link to O2C Data Model (for the O2C Discovery Accelerator, Process-specific)

https://docs.uipath.com/process-mining/docs/input-tables-of-the-order-to-cash-discovery-accelerator

High-level diagram of transformation steps



Remarks:

- This diagram shows only what happens in the transform step (2).
- The dotted-red line shows the difference between the data model for the Discovery accelerators (P2P/O2C) and TemplateOne.
- This diagram is a visual representation of the steps described in the Connector Development Kit in Appendix G.

Data Model compatibility notes.

Data Models for TemplateOne and for process specific DAs **both include events,** because event logs are necessary for Process Mining. The Data Model for TemplateOne is small and flat (one table). The Data Models for Process-specific DAs include entities (multiple tables). The initial DA data models (P2P, O2C) are not fully compatible with TemplateOne. However, since TemplateOne is small and simple and includes events it is possible to add entities to an TemplateOne data model. See Requirements below for how the SQL of the connector can be built and organized so that this "evolution' from a simple to a complex data model is made easier.

Requirements:

- Transformations should be written in ANSI SQL:1999 with the SQL:2003 analytic extensions.

- The SQL is built as SQL files for
    - Extract (name … _Extract)
    - Transform (name... Transform)
        - For TemplateOne Connectors, cleanly separate the logic that gets entities from the logic that generates the process mining event log. In particular, adhere to the guidelines of the Connector Development Kit in Appendix G
        - For TemplateOne Connectors, include as many relevant Tags as possible, at least 5. Prioritize those Tags that are of high value/importance within the process for which the Connector is built. See Chapter 6 of the Connector Development Kit in Appendix G.
        - For TemplateOne Connectors, include as many relevant Due Dates as possible, at least 5. Prioritize those Due Dates that are of high value/importance within the process for which the Connector is built. See Chapter 6 of the Connector Development Kit in Appendix G.
        - For TemplateOne Connectors, include work estimates for each activity. See Chapter 4 Section “Optional Attributes” of the Connector Development Kit in Appendix G.
        - For DA Connectors: adhere to the guidelines of chapters 1 – 4 of the Connector Development Kit in Appendix G
    - DBT Core project with the same sub-structure as the above
- The transformation can be run on Snowflake
- The transformation can be run on PostgreSQL
- The transformation can be run on SQL Server 2019
- The output of the transformation adheres to the defined Schema. This can be checked by running the validation queries that UiPath provides.
- All Validation queries pass.

Note: It is expected that more guidelines for writing connectors become available in the next few months.