Integrating a templated configurator into the ETL process

**Automated OMOP-CDM pipeline for the new EBMT Registry**

**Background:** The configurator is a web application solution that templates the process of converting the EBMT Registry’s complex data-entry forms base into OMOP-CDM. This solution allows for new configurations to be easily entered into the application through the user interface and applied to the Extract–Transform–Load (ETL).

The user enters the new configuration through the EBMT Registry’s application user interface. The new configuration is saved to the application database. The nightly pipeline copies the relevant patient data and the configurations to the analytical database (icopy process), and uses the configurations during the ETL to map the patient data to OMOP-CDM format (integration process), and stored on the analytical database. The process is then tested for quality assurance (test process).

**Methods**

1. Configuration data for each field that is recorded is stored in the Field table (1a). The Data Service Option table (1b) stores configuration data for fields with pre-fixed options. Patient data is stored in the Field Response table (1c).

   ![Table 1a](image1.png)

   The ETL process first builds an intermediary STEM table with all the necessary data for each field. Using the configurations in the Field table, the process captures the date of each data point, even when it is not included in the Field Response.

   ![Table 1b](image2.png)

   The ETL process then joins the STEM and Concept tables on concept_ids to retrieve the domain and populates the associated clinical table.

   ![Table 1c](image3.png)

   The ETL process is tested for quality assurance (itest process).

2. **Conclusion:** Templating mapping configurations via user interface keeps the OMOP-CDM at the forefront when expanding data collection and allows for seamless expansion of the application database without needing to update the code, pipeline, or ETL process. As a result, this pipeline can be maintained by any non-technical team member that is familiar with OMOP-CDM.