Development of an ETL Process for Bulk and Incremental Load of German Patient Data into OMOP CDM Using FHIR

Elisa Henke[1], Yuan Peng[1], Ines Reinecke[1], Michéle Zoch[1], Martin Sedlmayr[1]
[1] Institute of Medical Informatics and Biometry, Carl Gustav Carus Faculty of Medicine, Technische Universität Dresden, Germany

This work is part of the project MIRACUM, funded by the German Ministry of Education and Research (FKZ 01ZZ1801A/L).

Background

Motivation:

- The Use Case 'Alerting in Care – IT Support for Patient Recruitment' [1] in MIRACUM (Medical Informatics in Research and Care in University Medicine) aims to develop a Clinical Trials Recruitment Support System (CTRSS).
- This system suggests patients for clinical trials based on data in the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM).

Objective:

- To provide data for the CTRSS we need to design and develop an ETL (Extract-Transform-Load) process for filling OMOP CDM using Fast Healthcare Interoperability Resources (FHIR) profiles from MI-I and MIRACUM as data source.
- The ETL process has to support an initial (bulk) load as well as near real time or at least once a day updates (incremental load) of the data in OMOP CDM, to enable quick recruitment.

Methods

Semantic Mapping

FHIR Profiles
- Patient
- Encounter
- Condition

[2]

- Medication
- MedicationAdministration
- Observation
- Procedure

[3]

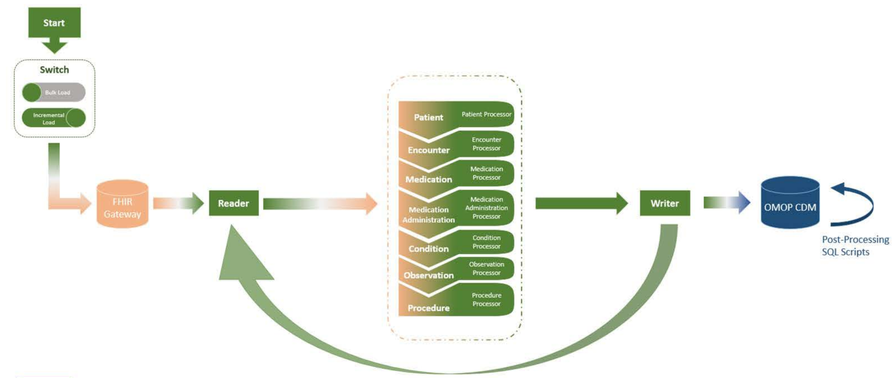
OMOP CDM tables [4]
- PERSON
- LOCATION
- DEATH
- VISIT_OCCURRENCE
- VISIT_DETAIL
- OBSERVATION_PERIOD
- CONDITION_OCCURRENCE
- FACT_RELATIONSHIP
- PROCEDURE_OCCURRENCE
- OBSERVATION
- MEASUREMENT
- DRUG_EXPOSURE

SpringBatch Framework [5]

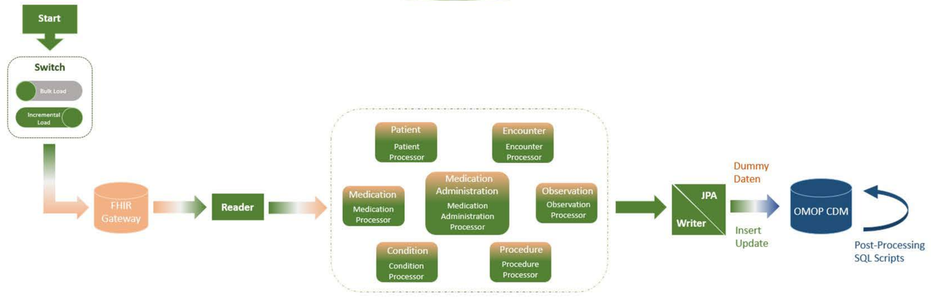
Results

We have designed and implemented an ETL process which transforms MI-I and MIRACUM FHIR resources to OMOP CDM. This ETL process uses a switch to select whether the ETL process is executed as bulk load or as incremental load.

bulk load
with predefined order
of processing
FHIR resources



incremental load
with random order
of processing
FHIR resources



Conclusions

- The developed ETL process can transform and load data from FHIR into OMOP CDM as bulk load or incremental load.
- Thus, patient data can be updated to enable rapid recruitment with the CTRSS based on OMOP CDM.
- In the future, it is our aim to:
  - use meta data from FHIR and OMOP CDM to automate the ETL process
  - update the ETL process to new versions of the FHIR profiles from MI-I

References:
[1] Reinecke I, Gulden C, Kümmel M, Nassirian A, Blasini R, Sedlmayr M. Design for a Modular Clinical Trial Recruitment Support System Based on FHIR and OMOP. Stud Health Technol Inform. 2020 16;270:158-162.
[2] Medical Informatics Initiative Germany. Basismodule des Kerndatensatzes der MII. Available from: https://www.medizininformatik-initiative.de/de/basismodule-deskerndatensatzes-der-mii.
[3] MIRACUM. MIRACUM Core Implementation Guide – Table of Contents. 2020. Available from: https://fhir.miracum.org/core/toc.html.
[4] Observational Health Data Sciences and Informatics. OMOP CDM v5.3.1. Available from: https://ohdsi.github.io/CommonDataModel/cdm531.html.
[5] Ward L, Syer D, Risberg T, Kasanicky R, Garrette D, Lund W, Minella M, Schaefer C, Hillert G, Renfro G, Bryant J, Hassine M B. Spring Batch – Reference Documentation. 2021. Available from: https://docs.spring.io/spring-batch/docs/current/reference/html/index.html.

Contact: Elisa.Henke@ukdd.de