

Introduction

Realizing our vision to enhance interoperability, connectivity and standardization of disparate healthcare data, EvidNet is revolutionizing the way medical researchers generate evidence translatable to clinical practice through the establishment of a fully interoperable, seamless, real-time integrated healthcare data network, ultimately benefiting health professionals and patients. As one of our major milestones, we are participating in a South Korean government project¹ aiming to convert 54M patient data from 43 hospitals within three years and have so far successfully converted data from 9 hospitals into the OMOP CDM. In this poster, we share in detail our experiences in implementing the OMOP CDM and OMOP ecosystem in South Korea.

ETL Process

Efficient management of time and manpower is crucial to convert the clinical data of 43 hospitals into CDM within three years. However, as different EMR systems are used in different hospitals, it is heavily time-consuming to comprehend the data structure and terminology of every hospital, individually design ETL processes and then perform serial conversions. To solve this problem, we have identified phases in the ETL process that can be standardized and developed an automated conversion tool. Specifically, we have designed a common interface table that comprises all the respective source data necessary for conversion into CDM and automated common tasks such as application of conversion rules and terminology mapping based on this interface table. A typical CDM conversion and data quality management process is illustrated in **Figure 1**.

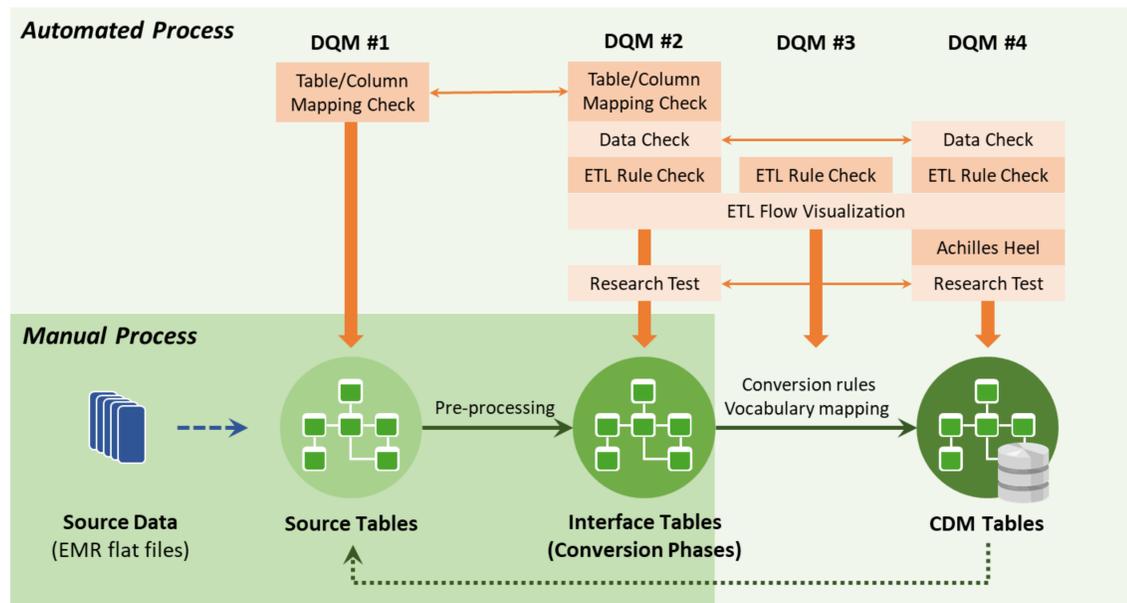


Figure 1 CDM conversion and data quality management process

Three main advantages of our automated conversion tool are as follows: (a) complete an ETL process in a short period of time with consideration of data quality, (b) enable parallel CDM conversions and periodic maintenance operations, (c) add new features and processes (extensibility).

As the first extension, we are currently developing multiple features that will enable review of data quality in parallel to the conversion process. This extension will enable the detection of successful application of conversion rules, identification of abnormal values generated during the conversion process, review the final CDM data through the execution of Achilles Heel rules, as well as recommend appropriate solutions upon error occurrences, ultimately improving the experience of the ETL task and quality of the final CDM data.

Terminology Mapping

Standard codes commonly used by the South Korean national health insurance system and healthcare organizations such as EDI (electronic data interchange) codes and ICD-10-based KCD-7 disease classification codes, along with local codes used by individual hospitals are continuously increasing. To respond to these changes rapidly, EvidNet has set up a systematic mapping procedure with medical terminology experts and developed a mapping management software capable of stably mapping up to 3,000 concepts per month. In addition, a mapping feature using NLP (Natural Language Processing) algorithms will be implemented into the mapping management software in order to increase the accuracy and speed of the mapping process. A typical terminology mapping process is illustrated in **Figure 2**.

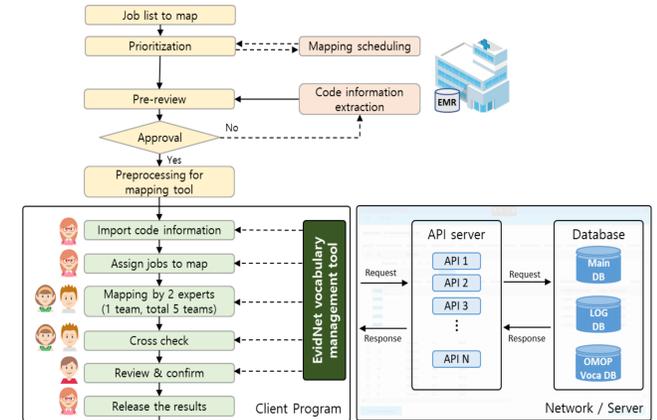


Figure 2 Terminology mapping process

Coordinating Platform

EvidNet is in active development of a web-based coordinating platform FEEDER-NET (Federated E-health Big Data for Evidence Renovation Network, **Figure 3**), enabling the management and deployment of a CDM-based distributed research network. Once live, FEEDER-NET will serve as an innovation sandbox for data-driven services, building a sustainable virtuous cycle ecosystem between data owners, service providers and data users. Key features of FEEDER-NET are listed in **Table 2**.

Feature	Details
OHDSI Tools	- Centralized ATLAS - Data visualization through embedded ACHILLES
CDM Extensions	- Genomic, GIS, PGHD, radiology, national claim data (HIRA-NPS), national ER registry (NEDIS), etc. integrated
Multicenter Research	- Register and monitor IRB status - Manage and track data access and analyses - Invite research partners within network
Vertical Services	- Open innovation ecosystem with multiple vertical services - Open APIs available for vertical service providers

Table 2 Key features of FEEDER-NET coordinating platform

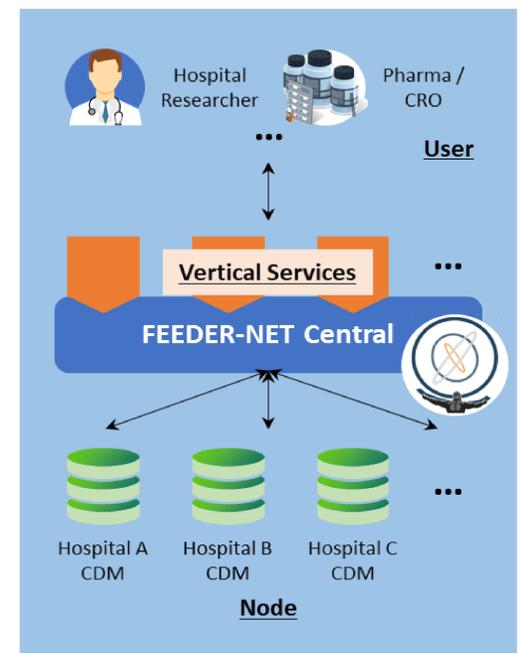


Figure 3 FEEDER-NET coordinating platform

Conclusion

Honoring OHDSI's philosophy, EvidNet is actively involved in spreading the OMOP CDM throughout South Korea and we believe our experiences will contribute to developing best practices in implementing the OMOP CDM and OMOP ecosystem on a nationwide scale. Through our technical expertise and coordinating platform, we strive to accelerate innovative medical researches and groundbreaking healthcare business opportunities.

References

1. SW Lee, SC You, RW Park. Development of FEEDER-NET (Healthcare Big Data Ecosystem) based on OMOP-CDM in Korea. 2018 OHDSI Symposium poster abstract.