

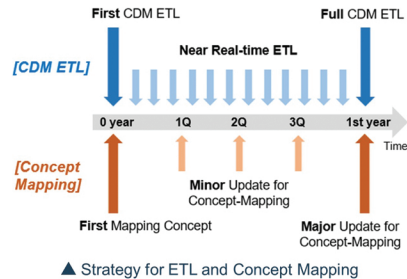
Introduction

- This study introduces the near real-time incremental OMOP-CDM ETL (hereafter, NRT ETL) system.
- This NRT ETL system will broaden the research scope of OMOP-CDM to the studies requiring up-to-date data, such as COVID-19 studies or prospective pragmatic clinical trials studies.

Strategies for NRT ETL

- We established the following NRT ETL strategies in advance.

Items	Desc.
ETL cycle	• Daily or weekly or monthly
Data scope	• Incremented data since the last (full / near real-time) ETL
Data sync. strategy	• Full ETL in a year cycle + near real-time ETL
Concept mapping strategy	• Major mapping in a year cycle + minor mapping in a quarter cycle
Backup strategy	• Archiving the latest three CDM versions



NRT ETL System Development

ETL System Architecture

- The NRT ETL system was developed using open-source softwares such as Digdag, Embulk, Spark, Drill, Parquet, which enhance system scalability.
- From original data to final CDM data, four intermediate staging files were produced:
 - 1) **ODS**: original loaded data
 - 2) **Interface**: column mapped data
 - 3) **Intermediate**: meta code mapped data
 - 4) **CDM**: medical code mapped data

NRT ETL System

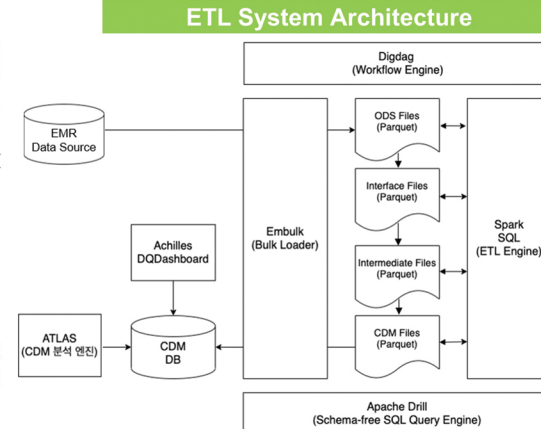
- We improved the ETL system to support NRT ETL by sophisticating the following modules.

1) Extraction of original data

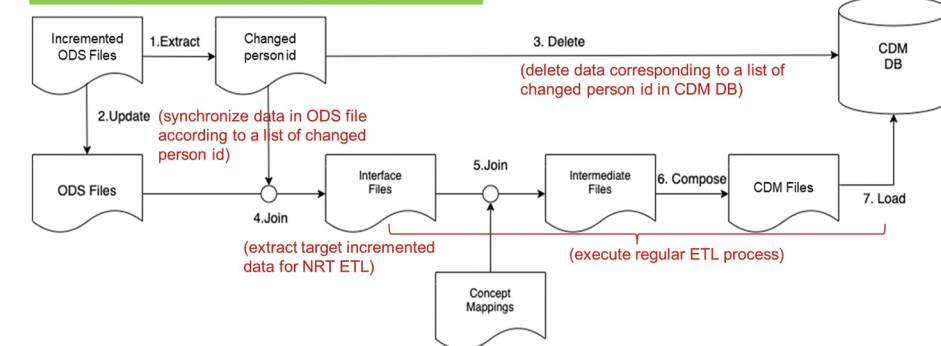
: Using change data capture (CDC) or [BETWEEN] query for the [DATE] column, we captured incremented data

2) Synchronization of [person_id]

: With a derived list of changed person_id, we synchronized ODS file and CDM DB, and executed the ETL process



Workflow for NRT ETL



Results & Conclusion

- We developed the NRT ETL system using Ajou university hospital CDM, which has an Oracle CDC software.
- Currently, NRT ETL system was successfully applied in 11 hospitals (in green ba
- ckground) and is under development in 12 hospitals in South Korea as of

Hospitals of NRT ETL System	
Chonnam National Univ. Hospital *	Chonnam National Univ. Hwasun Hospital *
Daegu Catholic Univ. Medical Center *	Dankook Univ. Hospital *
Gangwon National Univ. Hospital *	Kungpook National Univ. Chilgok Hospital *
Kyung Hee Univ. Hospital at Gangdong *	Kyungpook National Univ. Hospital *
Myongji Hospital *	Wonkwang Univ. Hospital *
Wonju Severance Christian Hospital *	Ajou Univ. Hospital
Bucheon Sejong Hospital	Ewha Womans Univ. Mokdong Hospital
Ewha Womans Univ. Seoul Hospital	Incheon Sejong Hospital
Kangbuk Samsung Hospital	Kyung Hee Univ. Medical Center
National Health Insurance Service Ilsan Hospital	Soon Chun Hyang Univ. Hospital Bucheon
Soon Chun Hyang Univ. Hospital Cheonan	Soon Chun Hyang Univ. Hospital Gumi
Soon Chun Hyang Univ. Hospital Seoul	

Acknowledgement

- Acknowledgment : This research was funded by the Bio Industrial Strategic Technology Development Program (20003883, 20005021) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea) and a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HR16C0001).
- Conflict of interest : W. Lee, S. Lee, and J. Y. Kim are employees of EvidNet.
- Contact : veritas@ajou.ac.kr