

# Transformation heterogener intensivmedizinischer Datenbanken in das OMOP-CDM

(Bachelorarbeit)



ANNA VAABEN

## Motivation

Diese Bachelorarbeit wird im Rahmen des ASIC-Use-Case des SMITH-Projekts, welches Teil der Medizininformatik-Initiative ist, geschrieben. Bei der intensivmedizinischen Behandlung werden zur Kontrolle des Gesundheitszustandes des Patienten viele relevanten Informationen, wie Vitalparameter gespeichert. Die Speicherung dieser Daten erfolgt im Regelfall in krankenhausinternen Datenbanken. Dabei sind im Allgemeinen unterschiedliche Datenmodelle im Einsatz. Diese Diversität erschwert eine Analyse der Daten, da für jeden Datensatz ein auf das benutzte Schema angepasste Methode benötigt wird. Mit Hilfe eines einheitlichen Datensatzes mit einem definierten Datenmodell kann die Analyse und weitere Benutzung der Daten vereinfacht werden.

## Stand der Technik

Um Datenmodelle zu vereinheitlichen, werden Extract-Transform-Load-Prozesse (ETL) verwendet. Das Ziel hierbei ist es die Daten aus mehreren verschiedenen aufgebauten Datenbanken in eine einheitliche Zieldatenbank zu transferieren. Für die Struktur der Zieldatenbank gibt es unterschiedliche Standards. Für diese Arbeit wurde das OMOP CDM von der Organisation „Observational Health Data Sciences and Informatics“ kurz OHDSI gewählt. Dieser Standard hat sich bereits bei vielen internationalen Organisationen etabliert. Desweiteren gibt es verschiedene Tools, die von OHDSI zur Verfügung gestellt werden. Diese können beim Transformationsprozess der Datenschemata unterstützen und für die Analyse und Veranschaulichung von Daten im OMOP-CDM Schema angewendet werden. Es existiert Literatur, die sich mit der Übertragung von verschiedenen Datenbanken zu OMOP CDM befasst. Ungeachtet dessen ist eine individuelle Betrachtung der Datenbanken für eine Konvertierung erforderlich, da die Datenbanken in ihren Schemata stark variieren.

## Zielsetzung

Ziel der Arbeit ist die Transformation der Daten von verschiedenen dem Projekt vorliegenden Datenbanken in den OMOP CDM Standard V5.3.1, um die Benutzbarkeit der intensivmedizinischen Patientendaten zu erhöhen. Darüber hinaus sollen die überführten Daten im Kontext des ASIC-Use-Cases analysiert werden. Dies soll mit Hilfe der von OHDSI bereitgestellten Analyse Tools realisiert werden. Die Analyse dient der Evaluation der Nutzbarkeit der neu transformierten Daten, sowie der Validierung der neu erstellten Datenbank.

## Geplante Vorgehensweise

Nach einer vorangegangenen Literaturrecherche erfolgt die Transformation der genannten Datenbanken in den OMOP CDM Standard. Hierbei wird auf die Hilfe der bereits vorhandenen Transformationstools zurück gegriffen. Um das neu entstandene Datenschema anschließend zu validieren, wird eine Genauigkeitsprüfung auf den transferierten Daten durchgeführt und mit einer analogen Analyse der Originaldaten verglichen. Darüberhinaus werden quantitative Statistikprüfungen der transformierten Daten durchgeführt und ebenfalls mit den statistischen Eigenschaften der Originaldaten verglichen.