

Gematik-Workshop „Konzepte der Datenintegration für die Sekundärnutzung“, 20.06.2024

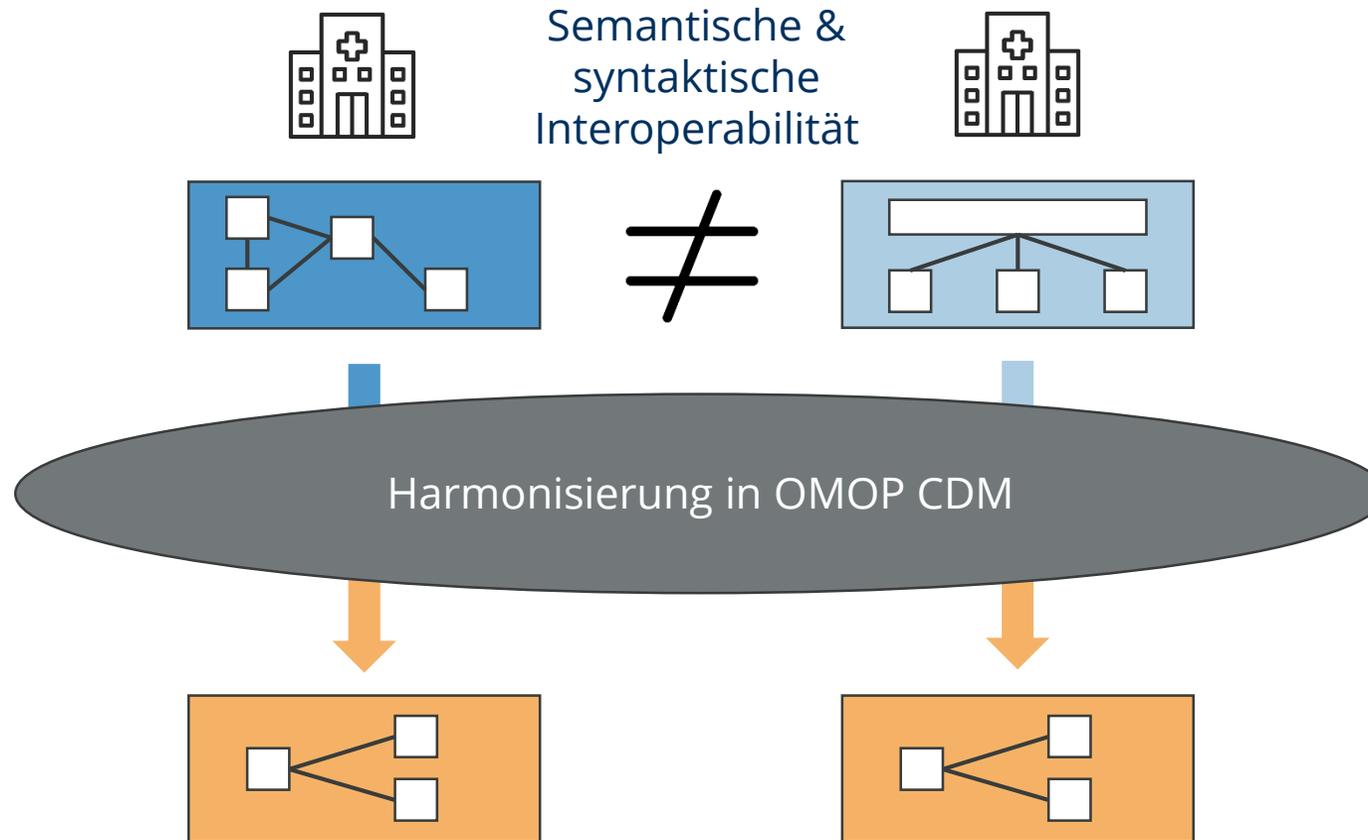
Harmonisierung klinischer Daten in OMOP CDM

Elisa Henke und Yuan Peng

Institut für Medizinische Informatik und Biometrie,

Medizinische Fakultät Carl Gustav Carus der Technischen Universität Dresden

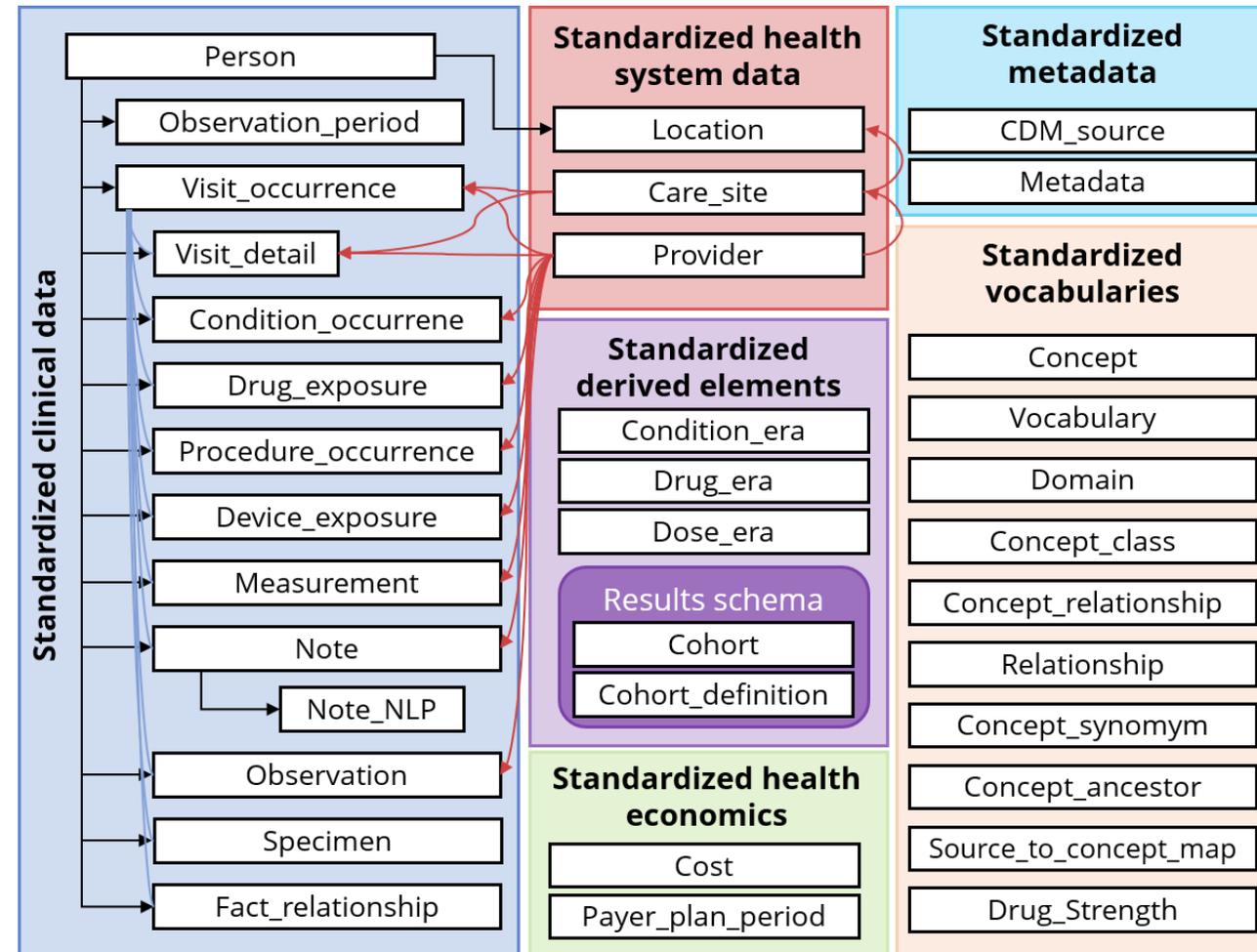
Offene Herausforderung



Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM)



- patient*innenzentriertes Datenmodell
- standardisierte Vokabulare und Datentabellen
- Verarbeitung von Daten aus verschiedenen Quellen
- bewahrt Datenprovenienz und Quelldaten



Generischer Datenharmonisierungsprozess für OMOP CDM

 **Quantitative data quality analysis**

 **Qualitative data quality analysis**

 **ETL-prozess**

 **Structural mapping**

 **Semantic mapping**



Dataset specification 

Data profiling 

Henke et al. *BMC Medical Informatics and Decision Making* (2024) 24:58
<https://doi.org/10.1186/s12911-024-02458-7>

BMC Medical Informatics
and Decision Making

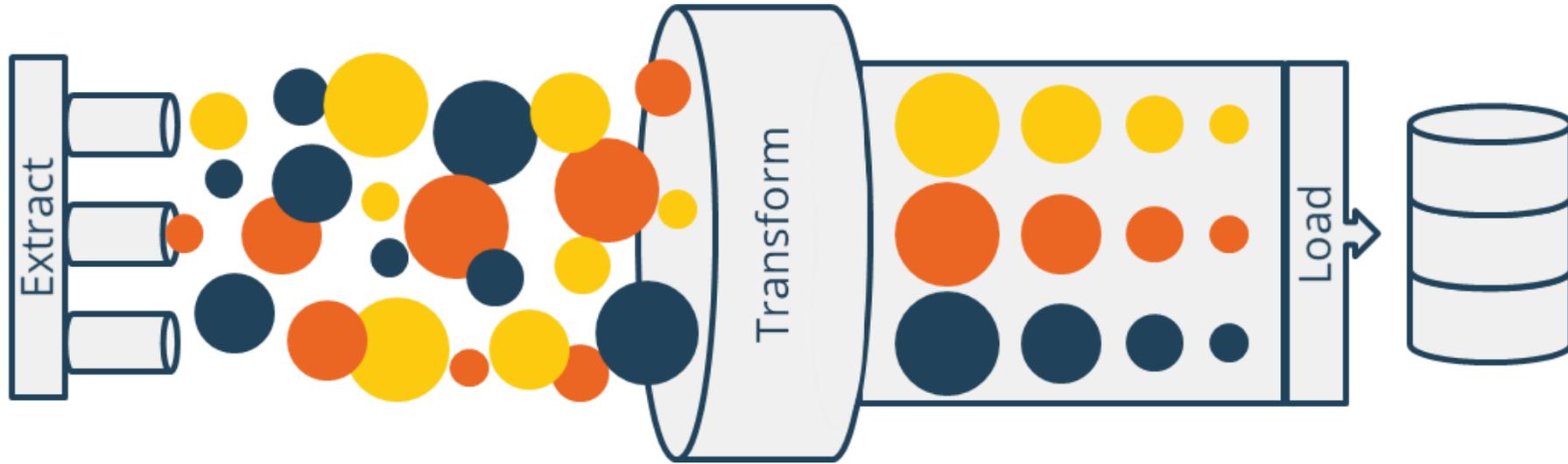
RESEARCH

Open Access

Conceptual design of a generic data harmonization process for OMOP common data model

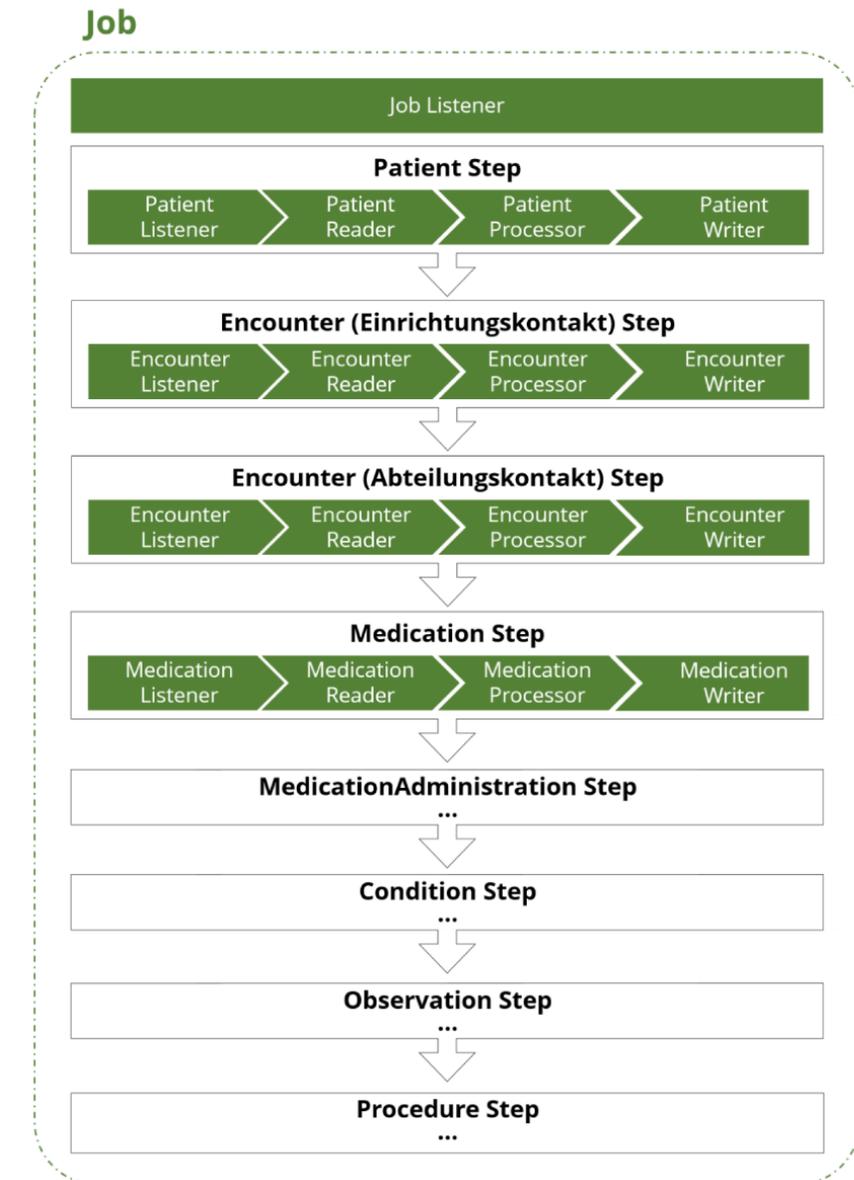
Elisa Henke^{1*}, Michele Zoch¹, Yuan Peng¹, Ines Reinecke², Martin Sedlmayr¹ and Franziska Bathelt³

ETL-Prozess



Architektur des FHIR-to-OMOP ETL-Prozesses

- **Job**
 - Verarbeitung der FHIR-Ressourcen nach OMOP CDM
 - **Job Listener**
 - Laden von neu erstellten semantischen Mappings nach OMOP CDM
 - **Step**
 - Repräsentation eines unabhängigen, sequentiellen Teils pro FHIR-Profil
 - **Step Listener**
 - Herunterladen benötigter Vokabulare aus OMOP CDM in den Arbeitsspeicher
- **Step Reader**
 - Lesen der FHIR-Ressourcen aus der Quelle
 - **Step Processor**
 - Transformation der eingelesenen FHIR-Ressourcen
 - **Step Writer**
 - Schreiben der transformierten FHIR-Ressourcen nach OMOP CDM



Herausforderungen



code	S	Σ	1..1	CodeableConcept
coding	S	Σ	1..*	Coding
icd10-gm	S	Σ	0..1	Coding Binding
alpha-id	S	Σ	0..1	Coding Binding
sct	S	Σ	0..1	Coding Binding
orphanet	S	Σ	0..1	Coding
text		Σ	0..1	string

```
{
  "code":{
    "coding":[
      {
        "system":"http://fhir.de/CodeSystem/bfarm/icd-10-gm",
        "version":"2020",
        "code":„I50.1",
        "display":Linksherzinsuffizienz"
      }
    ]
  }
}
```



code	S	Σ	1..1	CodeableConcept
------	---	---	------	-----------------

```
{
  "code":{
    "coding":[
      {
        "system":"http://hl7.org/fhir/sid/icd-9-cm",
        "version":"2020",
        "code": 428.1",
        "display":Left heart failure"
      }
    ]
  }
}
```

Probleme



MEDIZIN
INFORMATIK
INITIATIVE



Hoher Zeitaufwand



Schlechte Wartbarkeit



Schlechte Wiederverwendbarkeit



US Core

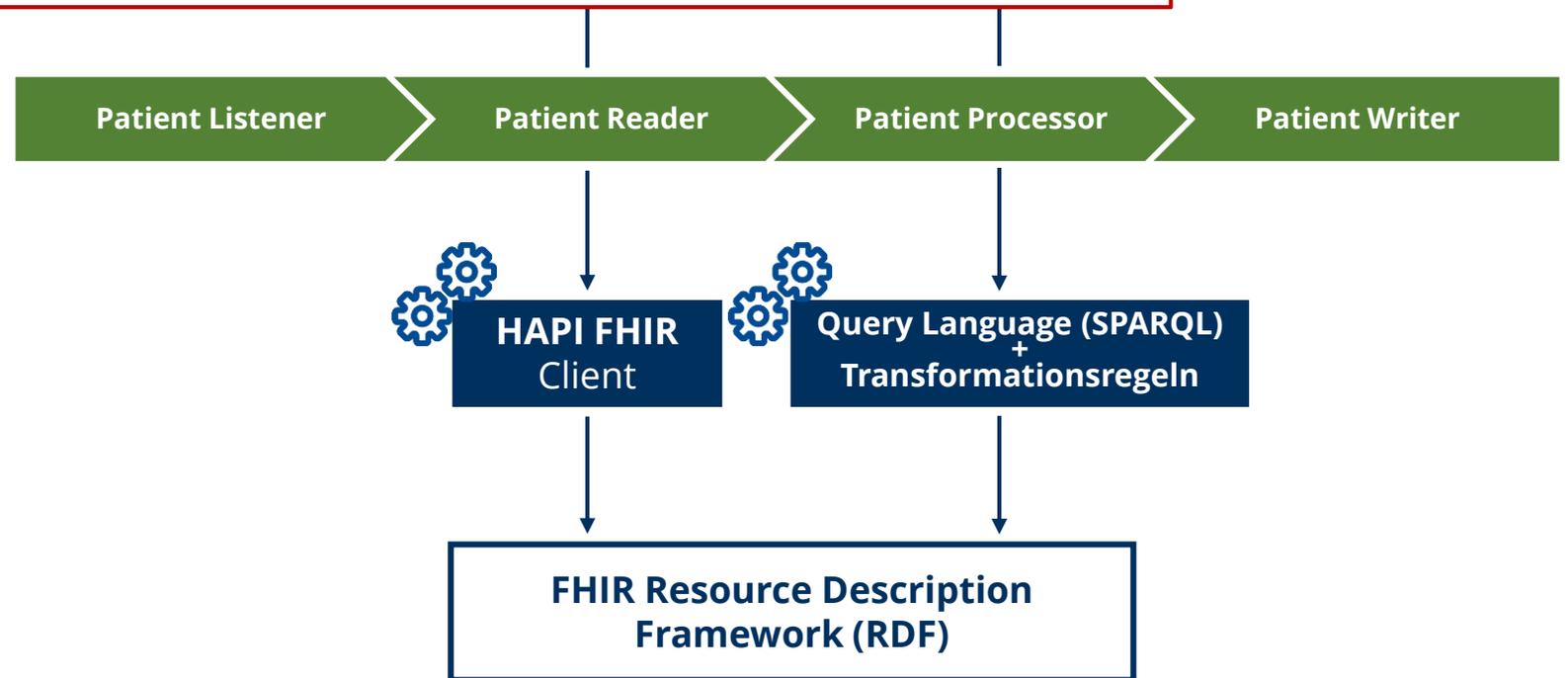


Metadaten-getriebener ETL-Prozess



Use of Metadata-Driven Approaches for Data Harmonization in the Medical Domain: Scoping Review

Yuan Peng¹ ; Franziska Bathelt² ; Richard Gebler¹ ; Robert Gött³ ; Andreas Heidenreich⁴ ; Elisa Henke¹ ; Dennis Kadioglu^{4, 5} ; Stephan Lorenz¹ ; Abishaa Vengadeswaran⁵ ; Martin Sedlmayr¹



Metadaten-getriebener FHIR-to-OMOP

Zusammenfassung

- wiederverwendbarer ETL-Prozess zur Überführung des MII Kerndatensatz nach OMOP CDM
- Befähigung aller Datenintegrationszentren zur Überführung von Daten nach OMOP CDM
- kontinuierliche Weiterentwicklung des ETL-Prozesses durch das Dresdner ETL-Team
- Weitere Informationen:
 - GitLab-Repository „FHIR-to-OMOP“ inklusive Dokumentation der Mappings
 - Peng et al., 2023. *An ETL-process design for data harmonization to participate in international research with German real-world data based on FHIR and OMOP CDM*
 - Henke et al., 2023. *An Extract-Transform-Load Process Design for the Incremental Loading of German Real-World Data Based on FHIR and OMOP CDM: Algorithm Development and Validation*



Vielen Dank für Ihre Aufmerksamkeit.

 Quantitative data quality analysis

 Qualitative data quality analysis

 ETL-process

 Structural mapping

 Semantic mapping



Dataset specification 

Data profiling 

Vocabulary identification 

Coverage analysis of vocabularies 

Kontakt

Medizinische Fakultät Carl Gustav Carus
der Technischen Universität Dresden

Fetscherstr. 74 (Haus M161)

01307 Dresden

E-Mail: Elisa.Henke@ukdd.de

Yuan.Peng@ukdd.de

