Mapping from openEHR to FHIR and OMOP CDM to support interoperability for infection control

Mapping von openEHR zu FHIR und OMOP CDM zur Unterstützung der Interoperabilität für die Infektionskontrolle

Abstract

Healthcare delivery and health outcomes of patients could significantly improve if the different electronic health systems of medical institutions were interoperable, that is if they could exchange and combine their data in a meaningful way. In addition, patients are increasingly becoming active sources of data with the growing use of health apps. This valuable information should not be lost, but rather be integrated with the patients' data in an open platform approach where the distance between caregiver and patient is reduced. Charité Universitätsmedizin Berlin and Berlin Institute of Health (BIH) are members of HiGHmed, a German consortium where eight university hospitals have agreed to the cross-institutional data exchange through novel medical informatics solutions. In this paper, we describe our approach to improve interoperability for the use case Infection Control of the HiGHmed project. Starting from the openEHR standard, we performed a syntactic mapping to the recommended standard *Fast Healthcare Interoperability Resources (FHIR)* and to the *Observational Medical Outcomes Partnership Common Data Model (OMOP CDM)*. FHIR enables fast exchange of data thanks to the discrete data elements into which information is organized, and OMOP CDM offers a data model specific for population-level analyses. Mapping is essential to identify which data have their correspondent elements in a different data model and can thus be converted into another standard without loss of information. As expected, not all openEHR and FHIR information could be mapped to OMOP CDM as many patient-level data are not important for population-level analysis. However, overall, mapping for the analyzed dataset was performed without major issues and we believe the use case will be able to exploit the advantages of the selected standards.

Keywords: interoperability, standard, mapping, data exchange, infection control, FHIR, OMOP CDM, openEHR

Eugenia Rinaldi[1]
Sylvia Thun[2,3]

1 Charité – Universitätsmedizin, Berlin, Germany

2 Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Germany

3 Hochschule Niederrhein, Informations- und Kommunikationstechnologien im Gesundheitswesen, Krefeld, Germany

Zusammenfassung

Die Gesundheitsversorgung und die gesundheitlichen Ergebnisse für Patienten könnten erheblich verbessert werden, wenn die verschiedenen elektronischen Gesundheitssysteme medizinischer Einrichtungen interoperabel wären, das heißt, wenn sie ihre Daten sinnvoll austauschen und kombinieren könnten. Darüber hinaus werden Patienten mit der steigenden Verbreitung von Gesundheits-Apps zunehmend selbst zu aktiven Datenquellen. Diese wertvollen Informationen sollten nicht verloren gehen, sondern in einem offenen Plattformansatz, bei dem die Distanz zwischen Leistungserbringer und Patient verringert wird, mit den Patientendaten integriert werden. Die Charité Universitätsmedizin Berlin und das Berlin Institute of Health (BIH) sind Mitglieder von HiGHmed, einem deutschen Konsortium, in dem sich acht Universitätskliniken auf den institutsübergreifenden Datenaustausch durch neuartige medizinische Informatiklösungen verständigt haben. In diesem

Beitrag beschreiben wir unseren Ansatz zur Verbesserung der Interoperabilität für den Anwendungsfall Infektionsschutz des HiGHmed-Konsortiums. Ausgehend vom openEHR- Standard führten wir ein syntaktisches Mapping zum empfohlenen Standard *Fast Healthcare Interoperability Resources* (FHIR) und zum *Observational Medical Outcomes Partnership Common Data Model* (OMOP CDM) durch. FHIR ermöglicht einen schnellen Datenaustausch dank der diskreten Datenelemente, in denen Informationen organisiert sind; OMOP CDM bietet ein Datenmodell, das speziell für Analysen auf Populationsebene geeignet ist. Das Mapping ist unerlässlich, um festzustellen, welche Daten ihre korrespondierenden Elemente in einem anderen Datenmodell haben und somit ohne Informationsverlust in einen anderen Standard umgewandelt werden können. Wie erwartet konnten nicht alle openEHR- und FHIR-Informationen auf OMOP CDM abgebildet werden, da viele Daten auf Patientenebene für Analysen auf Populationsebene nicht notwendig sind. Insgesamt jedoch wurde das Mapping für den analysierten Datensatz ohne größere Probleme durchgeführt. Der Anwendungsfall kann daher voraussichtlich die Vorteile der ausgewählten Standards nutzen.

Schlüsselwörter: Interoperabilität, Standard, Mapping, Datenaustausch, Infektionsschutz, FHIR, OMOP CDM, openEHR

Introduction

Under the umbrella of the Medical Informatics (MI) Initiative (https://www.medizininformatik-initiative.de), the German Ministry of Education and Research has funded the HiGHmed consortium with the aim of enhancing the efficiency of clinical research and improve patient care through novel medical informatics solutions and cross-institutional data exchange [1]. In particular, Charité and Berlin Institute of Health (BIH), together with seven other German universities that are members of HiGHmed, will be involved in the use case Infection Control whose primary aim is to merge all necessary pathogen-related data and information to establish a smart infection control system (SmICS).

While infection surveillance is already in place in hospitals and at national scale, it frequently suffers from a lack of data standardization and, consequently, from limited data integration, and limited availability of relevant data [2]. Research on multidrug-resistant organisms and their routes of transmission is based on analysis of information. The more information is available, the more precise analyses can be, and thus the more can clinical research benefit. For this reason, it is of extreme importance that hospitals and medical centers are enabled to combine and exchange their data.

To merge information coming from the different medical institutions of the HiGHmed consortium, standards for interoperability need to be adopted. Information has to be structured according to a common information model, and standard terminologies systems need to be employed. This enables smooth exchange of information and meaningful analysis of data. Regardless of the local systems used to store patient information in hospitals, standards ensure a level of interoperability that all hospitals can interface with. openEHR and FHIR are the most robust and complete healthcare data persistence and exchange specifications that support full semantic interoperability [3]. Both standards model the clinical and administrative data based on reusable patterns that describe the medical information. These patterns are called "Archetypes" in openEHR and "Resources" in FHIR. Archetypes are maximal data sets for a given single clinical concept and are expected to contain all the clinical information [4]. Archetypes are designed to be used in "Templates" that define specific use cases. In FHIR, resources only contain the most commonly used clinical information but can be extended with new data fields to support specific requirements.

The HiGHmed platform consists of an IHE XDS Affinity domain combined with an openEHR clinical repository. Consequently, also the information structure for SmICS will be based on openEHR. Archetypes and templates are accordingly being developed to model the infection control information. In comparison to FHIR resources, the extent and depth of information covered by archetypes brings a much higher level of complexity. The reason is that in openEHR, information is organized in a hierarchical structure where archetypes are processed together even when the target information is only in "child" Archetype. In FHIR, information is deconstructed into discrete data elements (resources) that can be processed individually; this approach makes it easier to transmit only the needed pieces of information without processing other unnecessary data. The tree-like structure of FHIR resources, with many recursive references, ensures a fast data exchange of independent resources. These features also make FHIR a particularly suitable standard to support the adoption of personal health record (PHR) model, with the patients' health records under their own control through the use of health mobile apps [5]. The government body of the Medical Informatics Initiative (MII) has recommended the use of the standard FHIR to all its participating consortia. As a MII consortium, HiGHmed needs to comply with its

recommendations and thus be in the position to exchange data via FHIR with the other MII members.

FHIR provides efficiency in information exchange, allowing access to granular patient health data along with cross references to other related information. As a consequence of the optimization for data exchange, the data format is not designed for storage and analysis like in traditional relational databases. In fact, FHIR resources are highly denormalized, with information grouped together, so that granular exchanges are fairly stand-alone [6]. On the other hand, such structure, optimized for data exchange, can be challenging when it comes to its use for data analytics purposes. The nested structure of information in openEHR is also denormalized and not optimized for exploration and analysis of data.

The OMOP common data model on the contrary offers optimized access to information for the sharing of health research data using relational databases [7]. OMOP is adopted by the Observational Health Data Sciences and Informatics OHDSI, a worldwide non-profit research alliance that focuses on open-source solutions for medical big data analysis.

Within MIRACUM, a consortium of the MI Initiative, a combination of OMOP CDM, FHIR and standardized vocabularies was used to develop a successful prototypical platform to perform statistical analysis and deploy resulting clinical decision models [8]. For their analyses, most institutions are used to relational databases which do not use tree-like formats with nested information. Therefore, to explore the possibilities for research analysis for our use case, we have considered the use of OMOP CDM.

Within the activities of the HiGHmed Interoperability Work Package, we have mapped archetypes to resources, and resources to OMOP tables. OMOP CDM facilitates research using data organization optimized for data analysis and predictive modeling. Moreover, OMOP provides a data model that uses standard terminology systems such as the Systematized Nomenclature of Medicine (SNOMED) and the Logical Observation Identifiers Names and Codes (LOINC) [9]. The use of standard vocabularies is also supported by FHIR and openEHR, and is very important for achieving semantic interoperability.

Finally, we performed an internal quality assessment of our work using the principles of mapping provided by the International Organization for Standardization (ISO) as reference [10].

Methods

For the use case "Infection Control", the HiGHmed consortium members have agreed upon a minimal dataset. The dataset contains the most relevant information that should be exchanged among institutions concerning a case of infection. It contains a selection of administrative, movement and microbiology data of patients.

The openEHR modelling group within the consortium has modelled the data using existing and new archetypes.

Local modifications of international archetypes were required to adapt the archetypes to the health care culture and the definition of health-related concepts in Germany [11]. Additionally, the microbiology-related archetypes were combined in an openEHR Microbiology Finding Template. The template was made available to the Berlin working group via the Clinical Knowledge Manager, the web platform for collaborative development, management and publishing of openEHR assets.

Starting from the agreed infection control minimum data set and the selected microbiology archetypes, we decided to proceed by steps in order to map the information of the use case to FHIR resources.

At first, we modelled the dataset in ART DECOR, an open-source tool and methodology that is commonly used to design and publish HL7 V3 templates of national (e.g. the Austrian electronic health record ELGA) and international EHR initiatives [12]. ART DECOR was useful for understanding the data model, for structuring the information and organizing it in a hierarchical view and associate pertinent terminologies (Figure 1).

Once we had a clear overview of the data structure, the dataset was mapped to the HL7 v2 messaging standard. HL7 v2 is the most widely implemented standard for healthcare for electronic data exchange in the clinical domain and it is currently used at Charité to exchange laboratory data [13]. Every element of the dataset was mapped to HL7 v2 using the support of Caristix (http://caristix.com/), a free web tool that supports HL7 interfacing. The HL7 theoretical mapping was then reviewed in comparison to its application in a real environment like Charité. It was helpful to interact with colleagues from Charité's communication server working group who were able to tell us how the HL7 v2 specifications had been implemented locally. They provided us with the information on how specific HL7 v2 segments were being received from the laboratory and used by the Charité Institute of Hygiene and Environmental Medicine. This step was important because it allowed us to have a close understanding of how information about laboratory findings is received and managed at Charité. HL7 v2 being a precursor of FHIR, this mapping activity was useful also to steer the information model in the desired direction.

From HL7 v2 to FHIR v4 the mapping was supported by the FHIR specifications that provide starting points for consideration [14]. All the openEHR and HL7 v2 items were matched with the resource elements using the information available on the FHIR website pages (https://www.hl7.org/fhir/resourcelist.html) that also offer mapping suggestions between HL7 v2 and FHIR.

Mapping was performed between openEHR archetypes and FHIR resources down to the data entry level to facilitate seamless use of both standards. Table 1 shows an example of the mapping performed on the pathogenic organism data.

Mapping from FHIR v4 to OMOP CDM v6 can be challenging because FHIR resources usually contain more information. We decided to ignore those elements that could not be mapped, mainly laboratory workflow details.

Figure 1: Art decor screenshot

Table 1: Mapping example between archetype and resource elements

Data	HL7 v2	openEHR	FHIR
Organism presence	OBX.8	CLUSTER.Laboratory test analyte.items[at0024]	Observation.interpretation
Organism name	OBX.2, OBX.5	CLUSTER.Laboratory test analyte.items[at0001]	Observation.value
Organism subtype	OBX.2, OBX.5	CLUSTER.Erregerdetails*.v0.items[at0047]	Observation.value
Colony count	OBX.2, OBX.5, OBX.6	CLUSTER.Erregerdetails.v0.items[at0035]	Observation.value
Colony frequency	OBX.8	CLUSTER.Erregerdetails.v0.items[at0003]	Observation.interpretation
Virulence factor	OBX.2, OBX.5	CLUSTER.Erregerdetails.v0.items[at0016]	Observation.value
MRO class	OBX.2, OBX.5	CLUSTER.Erregerdetails.v0.items[at0018]	Observation.value

However, almost all information concerning FHIR observation and specimen resources could be mapped to the elements of the OMOP tables *Measurement*, *Observation* and *Specimen*. For example, in FHIR all the laboratory test codes to investigate the microorganisms are defined in the element *Observation. Code* which can be mapped to the field *Measurement_concept_id* of the OMOP *Measurement* table. The code/id itself is provided by the selected terminology system, for example LOINC, which both standards support.

Our OMOP mapping was based on the Data Access Framework (DAF) FHIR Implementation Guide which describes the cross mapping of the two standards (https://www.ohdsi.org/web/wiki/doku.php?id= projects:workgroups:mappings_between_ohdsi_cdm_ and_fhir). Figure 2 shows the different mapping steps with the main tools used.
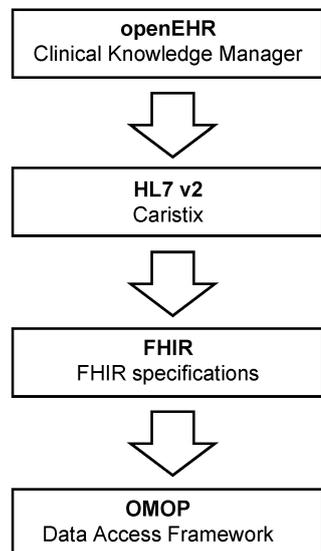
Figure 2: Mapping steps and main tools

Based on the Technical Report ISO TR 12300, which establishes measures to assess the quality and utility of a map, we performed an internal assessment of the quality of our mapping. This assessment is generally applied to terminology resources, but most of the principles stated there can also apply to syntactic mapping. This task was performed internally using the instructions and assessment scores provided by the ISO Report.

Results

For the use case infection control, we mapped the microbiology data from openEHR to FHIR. The data included information on the laboratory examination performed and the findings with pathogenic organisms and antibiogram details.

The openEHR archetypes

- *COMPOSITION.report-result*,
- *OBSERVATION.laboratory_test_ result*,
- *CLUSTER.specimen*,
- *CLUSTER.anatomical_location*,
- *CLUSTER.laboratory_test_analyte*,
- *CLUSTER.erregerdetails*

were mapped to the FHIR resources

- *ServiceRequest*,
- *Observation*,
- *Specimen*,
- *DiagnosticReport*

(Figure 3).

In particular, openEHR archetypes related to the laboratory results were mapped to the FHIR resource *Observation*. The archetypes concerning the specimen and the anatomical location were mapped to the FHIR resource *Specimen*.

As stated before, the archetypes are maximal data sets covering all aspects of a clinical concept. FHIR offers a minimum data set which could be extended to include more information in case of need. However, for the dataset analyzed, extensions were not necessary, the correspondence of elements between openEHR archetypes and FHIR resources for the microbiology data analyzed had a coverage of 100%. This was a very important result as inconsistencies between the models could pose a significant challenge for data interoperability [15].

Table 2 shows our own evaluation of the mapping performed according to the criteria of map quality included in the technical report ISO TR 12300. The first column of Table 2 shows the determinants of the map quality included in the ISO report; the second and the third columns show which score we assigned for each category to our mappings, openEHR to FHIR and FHIR to OMOP, respectively. The third column is shown as reference: it repre-
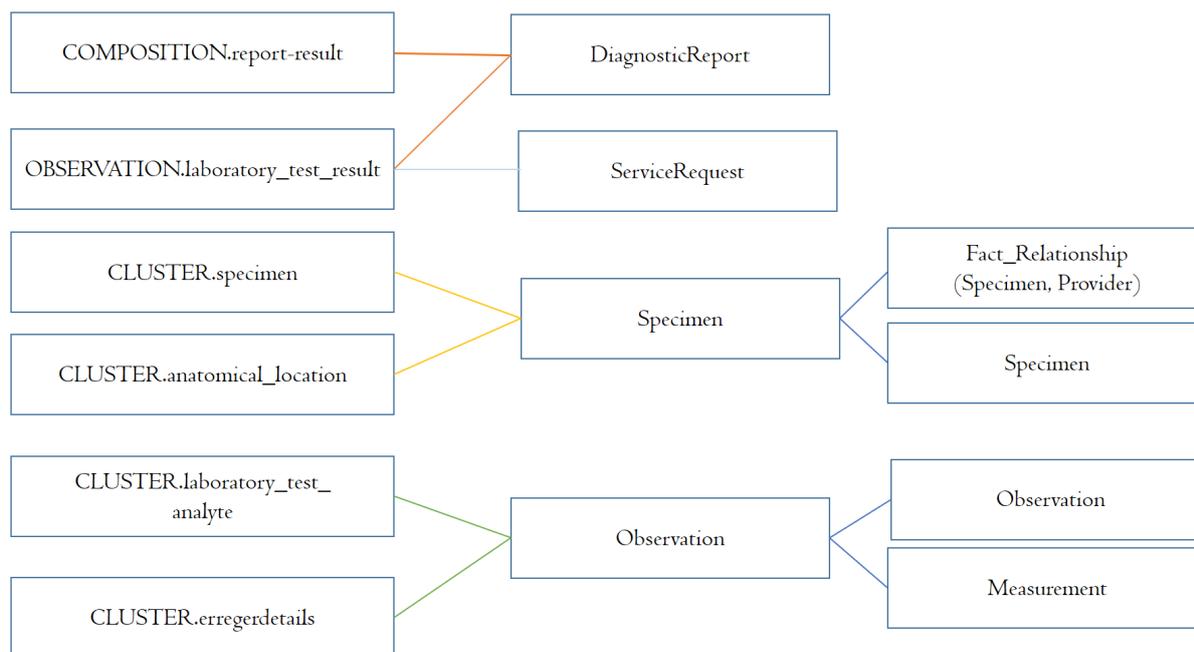
Figure 3: openEHR-FHIR-OMOP mapping

Table 2: Map quality assesment

Determinants of map quality	openEHR-FHIR	FHIR-OMOP	Suggested quality level for clinical use case
Common categorial structure	0	1	0
Shared semantic domain	0	0	1
Language and translation	N/A	N/A	0
Equivalence/Identification and publication	1	1	0
Equivalence assessment	0	0.9	0
Map set outliers	0	22%	0
Clear documentation of purpose	0	0	0
Currency of the map	0	1	0
Business arrangements	0	0	0
Methodology documentation	1	1	1
% of map validated	1	1	0
Method of validation	2	2	0
Decision making documentation	2	2	1
Tools	N/A	N/A	1
Skills	1	1	1
Map governance process documented	0	0	0
Map metadata versions and maintenance governance	1	1	0

sents the suggested quality level for a clinical use case according to the ISO report. The numerical scale is from 0 to 4, where 0 represents the perfect match or the highest score. The ISO report offers instructions on how to assign the score in each category. The determinants cover different aspects of the process of mapping, also including some organizational ones. "Equivalence assessment" is a particularly relevant parameter for our purpose as it represents a measure of the actual match between the mapped elements in two different standards. We assigned the value 0 to the mapping openEHR-FHIR where

we had perfect correspondence. We assigned a value of 0.9 to the mapping FHIR-openEHR as this was the average that resulted after scoring the equivalence to every single data element. Concerning the OMOP mapping, a good match (rated 2 and better) with FHIR resource elements was found for about 78% of the microbiology data set analyzed. The FHIR *Specimen* resource could be mapped to the OMOP table *Specimen*, and via the table *Fact_Relationship* to other tables, while the *Observation* resource could be mapped to the tables *Measurement* and *Observation*. All commentary fields are in OMOP

stored in a separate *Note* table. *DiagnosticReport* and *ServiceRequest* elements which contain laboratory report and more organizational information did not have an actual correspondence in OMOP. This was expected, as OMOP focuses on research and population level analysis and is not designed to describe laboratory workflow details.

Lower scores, such as "method of validation" or "decision making documentation" that were both rated '2', indicate that the map would benefit from having more people validating its content and from well documenting all the decisions and rules involved in the mapping process.

Overall, the mapping of the analyzed dataset from openEHR to FHIR and OMOP was satisfactory to pursue syntactic interoperability while supporting efficient health data exchange and research capabilities.

The full map between openEHR, FHIR and OMOP CDM elements can be found in Attachment 1 together with the assigned equivalence scores for each data element. The map offers several points of consideration for other working groups intending to use FHIR and OMOP CDM starting from openEHR or HL7v2. It is however important to remark that different contexts, specific implementations or other standard versions can greatly influence the results of the mapping.

Discussion

For the use case "Infection Control" of the HiGHmed project, the syntactic mapping of the data between the standards openEHR, FHIR and OMOP proved to be feasible without particular issues. All the above standards support the use of standard terminologies to also enable semantic interoperability. Our use case can thus benefit from fast data exchange without data loss and still enjoy adequate analysis capabilities. However, the dataset analyzed was limited to the microbiology data of the use case, and there might be new challenges should we consider bigger datasets. When mapping to FHIR, it is always possible to create profiles or extensions to best fit the requirements. In the case of OMOP, data which is not included in the data model is lost or has to be managed separately. This also reflects the focus of the two standards. FHIR concentrates on the patient level and on the simplicity of the queries to fetch specific data, OMOP focuses on the population level and therefore its CDM does not include all the information available in FHIR. In the quality map assessment, this gap finds evidence in the equivalence score "0.9" as opposed to "0" of the openEHR-FHIR mapping and in the percentage of outlier values (elements with equivalence rated 3 and 4). The assessment of the quality of the map helps to understand mapping weaknesses and also where it can possibly be improved; however, despite the scoring instructions, it remains a subjective assessment. There are increasing efforts to cross-reference standards and also to try to make the mapping process automatic. LinkEHR (https://linkehr.veratech.es/) for example, is described

as a tool with the functionality to transform openEHR archetypes to FHIR R4 Observation resources; however, not enough literature is available yet to prove its validity. Also within the consortium MIRACUM, a mapping converter tool has been developed to transform FHIR to OMOP. However, relying completely on automatic tools is not advisable and further reviewing is always recommended. Such tools are also usually bound to specific versions and might become obsolete as standards evolve. Additionally, the diversity of implementation of standards in different contexts or in different institutions remains indeed a big challenge when mapping. Our "handcrafted" approach was focused on a deep understanding of the standards, their elements and their general main purposes.

Notes

Competing interests

The authors declare that they have no competing interests.

Attachments

Available from
https://www.egms.de/en/journals/mibe/2021-17/mibe000221.shtml
1. Attachment1_mibe000221.pdf (224 KB)
   Map between openEHR, FHIR and OMOP CDM elements

References

1. Medical Informatics Initiative Germany. HiGHmed Consortium. [last accessed 2019 Nov 15]. Available from: https://www.medizininformatik-initiative.de/en/konsortien/highmed

2. Haarbrandt B, Schreiweis B, Rey S, Sax U, Scheithauer S, Rienhoff O, Knaup-Gregori P, Bavendiek U, Dieterich C, Brors B, Kraus I, Thoms CM, Jäger D, Ellenrieder V, Bergh B, Yahyapour R, Eils R, Consortium H, Marschollek M. HiGHmed – An Open Platform Approach to Enhance Care and Research across Institutional Boundaries. Methods Inf Med. 2018 Jul;57(S 01):e66-e81. DOI: 10.3414/ME18-02-0002

3. Allwell-Brown E. A Comparative Analysis of HL7 FHIR and openEHR for Electronic Aggregation, Exchange and Reuse of Patient Data in Acute Care. [last accessed 2019 Nov 15]. Stockholm: Karolinska Institutet, Stockholm University; 2016. Available from: https://ki.se/sites/default/files/migrate/eneimi_allwell_brown_a_comparative.pdf

4. Bosca D, Moner D, Maldonado JA, Robles M. Combining Archetypes with Fast Health Interoperability Resources in Future-proof Health Information Systems. Stud Health Technol Inform. 2015;210:180-4.

5. Roehrs A, da Costa CA, Righi RDR, Rigo SJ, Wichman MH. Toward a Model for Personal Health Record Interoperability. IEEE J Biomed Health Inform. 2019 Mar;23(2):867-73. DOI: 10.1109/JBHI.2018.2836138

6. HL7 FHIR Release 4. Persistent Storage. [last accessed 2019 Nov 28]. Available from: https://www.hl7.org/fhir/storage.html

7.   Choi M, Starr R, Braunstein M, Duke J. OHDSI on FHIR Platform Development with OMOP CDM mapping to FHIR Resources [Poster]. In: 2016 OHDSI Symposium; 2016 Sep 23-24; Washington D.C., USA.

8.   Gruendner J, Schwachhofer T, Sippl P, Wolf N, Erpenbeck M, Gulden C, Kapsner LA, Zierk J, Mate S, Stürzl M, Croner R, Prokosch HU, Toddenroth D. KETOS: Clinical decision support and machine learning as a service – A training and deployment platform based on Docker, OMOP-CDM, and FHIR Web Services. PLoS One. 2019;14(10):e0223010. DOI: 10.1371/journal.pone.0223010

9.   Khalilia M, Choi M, Henderson A, Iyengar S, Braunstein M, Sun J. Clinical Predictive Modeling Development and Deployment through FHIR Web Services. AMIA Annu Symp Proc. 2015;2015:717-26.

10.   ISO/TR 12300:2014(en), Health informatics - Principles of mapping between terminological systems. [last accessed 2020 Apr 7]. Available from: https://www.iso.org/obp/ui/#iso:std:iso:tr:12300:ed-1:v1:en

11.   Wulff A, Sommer KK, Ballout S; HiGHmed ConsortiumHaarbrandt B, Gietzelt M. A Report on Archetype Modelling in a Nationwide Data Infrastructure Project. Stud Health Technol Inform. 2019;258:146-50.

12.   Ott S, Rinner C, Duftschmid G. Expressing Patient Selection Criteria Based on HL7 V3 Templates Within the Open-Source Tool ART-DECOR. Stud Health Technol Inform. 2019;260:226-33.

13.   HL7 International. Product Brief. [last accessed 2019 Nov 18]. Available from: https://www.hl7.org/implement/standards/product_brief.cfm?product_id=185

14.   HL7 FHIR Release 5 Preview #3. v2 Messaging. [last accessed 2019 Nov 28]. Available from: https://build.fhir.org/comparison-v2.html

15.   Topaz M, Seger DL, Goss F, Lai K, Slight SP, Lau JJ, Nandigam H, Zhou L. Standard Information Models for Representing Adverse Sensitivity Information in Clinical Documents. Methods Inf Med. 2016;55(2):151-7. DOI: 10.3414/ME15-01-0081

Corresponding author:

Eugenia Rinaldi

Charité – Universitätsmedizin, Charitéplatz 1, 10117 Berlin, Germany

eugenia.rinaldi@charite.de