MEDINFO 2023 — The Future Is Accessible
J. Bichel-Findlay et al. (Eds.)

doi:10.3233/SHTI230919

Bringing Communities Together: Mapping the Investigation-Study-Assay-Model (ISA) to Fast Healthcare Interoperability Resources (FHIR)

Sophie A. I. KLOPFENSTEIN[a,b,1], Julian SASS[a] , Carina N. VORISEK[a], Felix JORCZIK[a], Carsten Oliver SCHMIDT[c], Matthias LÖBE[d], Martin GOLEBIEWSKI[e], Haitham ABAZA[e] and Sylvia THUN[b]
[a] *Core Facility Digital Medicine and Interoperability, Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Germany*
[b] *Institute for Medical Informatics, Charité – Universitätsmedizin Berlin, Germany*
[c] *University Medicine of Greifswald*
[d] *Institute for Medical Informatics (IMISE), University of Leipzig*
[e] *Heidelberg Institute for Theoretical Studies*
ORCiD ID: S. A. I. Klopfenstein https://orcid.org/0000-0002-8470-2258, J. Sass https://orcid.org/0000-0002-2068-7765, C. N. Vorisek https://orcid.org/0000-0001-7499-8115, F. Jorczik n/a, C. O. Schmidt https://orcid.org/0000-0001-5266-9396, M. Löbe https://orcid.org/0000-0002-2344-0426, M. Golebiewski https://orcid.org/0000-0002-8683-7084, H. Abaza https://orcid.org/0000-0003-0155-1149, S. Thun https://orcid.org/0000-0002-3346-6806

Abstract. Adhering to FAIR principles (findability, accessibility, interoperability, reusability) ensures sustainability and reliable exchange of data and metadata. Research communities need common infrastructures and information models to collect, store, manage and work with data and metadata. The German initiative NFDI4Health created a metadata schema and an infrastructure integrating existing platforms based on different information models and standards. To ensure system compatibility and enhance data integration possibilities, we mapped the Investigation-Study-Assay (ISA) model to Fast Healthcare Interoperability Resources (FHIR). We present the mapping in FHIR logical models, a resulting FHIR resources' network and challenges that we encountered. Challenges mainly related to ISA's genericness, and to different structures and datatypes used in ISA and FHIR. Mapping ISA to FHIR is feasible but requires further analyses of example data and adaptations to better specify target FHIR elements, and enable possible automatized conversions from ISA to FHIR.

Keywords. Metadata standards, interoperability, FAIR, open data strategy, fast healthcare interoperability resources, HL7 FHIR, investigation-study-assay, ISA Model

---

[1] Corresponding Author: Sophie A. I. Klopfenstein, Core Facility Digital Medicine and Interoperability, Berlin Institute of Health at Charité , Berlin, Germany; email: sophie.klopfenstein@charite.de.

1. Introduction

Publishing data and metadata – in the following referred to as (meta-)data – of health (research) datasets is a common objective of the European Commission and the Federal Government of Germany, as stated in the proposal of the European Health Data Space [1] and in the Federal Government's Open Data Strategy [2]. Both highlight the need of common data infrastructures, and exchange of health (meta-)data fulfilling the FAIR principles (findability, accessibility, interoperability, reusability) [3].

The German initiative National Research Data Infrastructure for Personal Health Data (NFDI4Health) [4] aims to facilitate and encourage collaboration between health related research communities. NFDI4Health created an infrastructure based on different existing platforms and provides a metadata schema (MDS) and diverse services, to collect, harmonize, analyze, and enable (meta-) data to be FAIR. The MDS is compatible with different trials registries and models [5] and its mapping to FHIR was tested [6].

Health and life science research communities introduced diverse – partially domain-specific – (meta-) data standards, terminologies, minimum reporting guidelines, and information models. The Investigation, Study, Assay (ISA) Model and Fast Healthcare Interoperability Resources® (FHIR®) are two examples: The ISA Commons community created a metadata framework composed of the generic ISA model and of a suite of tools, to capture, curate, manage, share and reuse datasets in biosciences. The core model enables to collect metadata at the project level (investigation), the research unit (study) and the experiments and measurements conducted (assay) [7]. Metadata is available as Tab-separated values (Tab) and JSON file formats. In contrast, FHIR® is a health information technology model developed by Health Level 7. It is a modular standard composed of information blocks (resources) that specify the structure and content to be captured. Resources can be linked through references. Additionally, FHIR extensions permit to cover supplementary information, and profiles list further rules and constraints to be applied to the resources. FHIR hosts content in XML, JSON or Turtle format [8].

As the NFDI4Health' local data hubs currently use the SEEK research data management system that integrates the ISA model [9], and the Central Health Study Hub [10] intends to use FHIR, we conducted a mapping of the ISA Model to FHIR (ISA-to-FHIR) and present our findings.

2. Methods

The ISA Model and Serialization Specifications (Version 1.0 from 2016.10.28) [11] as well as example data [12] were analyzed. We created a diagram and FHIR logical models (LM) – representing data models without using FHIR core resources –, to provide an overview of the ISA JSON schemas and ease the mapping. We included all elements grouped as "properties" (except "additional properties = false") and added information on data types, formats, references, and descriptions, where available. Cardinalities were derived from the ISA JSON schemas based on the JSON schema specification [13]. The LM acted as mapping source, FHIR® Release 4 (R4; v4.0.1 generated on November 11, 2019) as the target [8]. We did not aim to recreate a one to one the ISA structure in FHIR: We prioritized to map ISA properties, that can be covered with standard FHIR elements, or that are necessary for referencing purposes. We discussed internally the ISA-to-FHIR mapping suggestions regarding their correctness and conciseness and solved mapping disagreements based on majority votes. We analyzed the mapping qualitatively.

3. Results

We provide an implementation guide (IG) containing a schematic overview of the ISA JSON schemas, the ISA FHIR LM, and mappings to FHIR [14]. The LM contain 135 ISA properties in total. Mapping to FHIR was possible for 102 properties. We considered extensions for 37 ISA properties. We propose a network of FHIR resources to cover the ISA JSON schemas (**Figure 1**). While ten ISA JSON schemas are equivalent to FHIR resources, the remaining ten are represented by FHIR data types, integrated in diverse FHIR resources. The content of the “comment”-schema e.g., can be depicted as a FHIR “annotation” data type. To use the proposed FHIR resources, eight additional items would need to be added when profiling, as they are required by FHIR (cardinality 1..1).

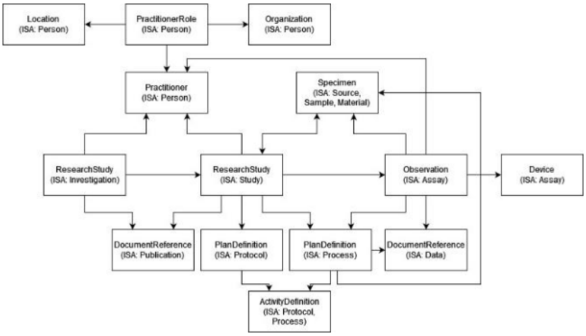

Figure 1. FHIR Information Model to Represent the ISA JSON Schemas.

During our mapping, we encountered challenges related to the:

- Detail of documentation: The level of detail of the specification documentation differed depending on the ISA format. Cardinalities and descriptions were sparse or missing in the ISA JSON specification; Information from ISA Tab could not always be used to guide the mapping, as its structure is different. We also discovered inconsistencies between the abstract model and specifications.
- Data types and cardinalities: ISA JSON includes a low number of data types that are generic, in contrast to many specific FHIR data types. Cardinalities were derived only from the ISA JSON schemas, and not from the requirement levels that are defined by RFC2119 and indicated in ISA JSON’s content rules. Cardinalities were not discussed with the ISA Community.
- Modelling approaches – Generic versus specific: ISA stores information in generic schemas that reference each other within a nested structure: For example, the “material” schema references to a second schema storing material “characteristic types”. A third schema stores the value associated with the material characteristic types and references back to the second schema. In FHIR, the concrete “material characteristic type” would be represented by the FHIR element name itself; values would be stored, e.g., as a Codeable Concept associated to this element. Depending on the concrete ISA data, different FHIR elements and resources might be suitable. In the IG, we provide ISA JSON examples related to material characteristics and mapped to FHIR.
- Representation of studies and related activities in FHIR: The proposed FHIR resources need adaptation to represent the study processes and the associated measurements. “Observation” might not fit in a study context, especially for

fundamental research: creating a new FHIR resource "Assay" could be a solution. Depending on the methods and procedures used, further resources such as "Imaging Study" or "Molecular Sequence" should be considered.

4. Discussion

To our knowledge, we presented the first ISA-to-FHIR mapping and analysis. Out of 135 ISA properties, 102 could be mapped to FHIR, but more than one third required extensions, leading to a deviation from FHIR's "80-20" rule: cover 80% of the common requirements with its core specification and more specific ones using extensions. Properties that were not mapped, were either not necessary in FHIR or could not be represented properly, e.g., because they were too generic for FHIR. ISA allows to store metadata from different communities and study types, but its structure impedes a precise mapping to FHIR. In our perception, FHIR allows a more unambiguous implementation in software through the use of profiles and IGs, while ISA provides more of a conceptual model. We tried to preserve ISA's original modularity to a certain extent, while aiming for a precise mapping: Some ISA schemas were not represented as single FHIR resources but rather integrated as elements or data types in other resources. Suitable standard FHIR elements and resources might already exist in R4 to cover ISA properties. However, selecting these requires clarifying open questions related to ISA's documentation, and to further analyze and process example data. Based on these activities, refining and automatizing the mapping, and profiling FHIR would be possible. While FHIR is already widely used in healthcare, its use in research is emerging [15]. Further refinements to Research Study, Observation, and definition resources – or even the creation of a new resource – would allow to better cover research activities and foster the use of FHIR in this context. Modifications are planned in FHIR R5 [16]; these might improve the current ISA-to-FHIR mapping and lower the number of extensions needed.

5. Conclusions

We showed that an ISA-to-FHIR mapping is mostly possible. It should be refined after clarification with the ISA commons group, and analyses of further example data. As NFDI4Health' MDS needs to comply with ISA and FHIR, we will integrate the findings from this study when updating its mapping to both models, and privilege reusable modules in FHIR, where meaningful. Additionally, profiling ISA in FHIR would be necessary to develop a conversion tool, ultimately fostering (meta-)data exchange and export between scientific domains such as biosciences, medicine, and related fields.

Acknowledgments and Competing Interests

This work was done as part of the NFDI4Health project (www.nfdi4health.de). We gratefully acknowledge the financial support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project Number 442326535. The generated content can be found on Simplifier.net - https://simplifier.net/guide/isa-to-fhir/home | Prof. Sylvia Thun is currently vice chair of HL7 Germany.

References

[1] European Commission, editor. Proposal for a Regulation of the European Parliament and of the Council on the European Health Data Space [Internet]. 2022 [cited 2022 Nov 1]. Available from: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52022PC0197

[2] Bundesministerium des Innern, für Bau und Heimat, editor. Open-Data-Strategie der Bundesregierung [Internet]. 2021 [cited 2022 Nov 12]. Available from: https://www.bundesregierung.de/resource/blob/997532/1940386/1d269a2ad1b6346fcf60663bdea9c9f8/2021-07-07-open-data-strategie-data.pdf?download=1

[3] Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016 Mar;3:160018, doi: 10.1038/sdata.2016.18.

[4] NFDI4Health. NFDI4Health National Research Data Infrastructure for Personal Health Data [Internet]. 2022. Available from: https://www.nfdi4health.de/en/

[5] NFDI4 Health Task Force COVID-19, Schmidt CO, Darms J, Shutsko A, Löbe M, Nagrani R, et al. Facilitating Study and Item Level Browsing for Clinical and Epidemiological COVID-19 Studies. In: Mantas J, Stoicu-Tivadar L, Chronaki C, Hasman A, Weber P, Gallos P, et al., editors. Studies in Health Technology and Informatics [Internet]. IOS Press; 2021 [cited 2022 Nov 30]. Available from: https://ebooks.iospress.nl/doi/10.3233/SHTI210284

[6] Klopfenstein SAI, Vorisek CN, Shutsko A, Lehne M, Sass J, Löbe M, et al. Fast Healthcare Interoperability Resources (FHIR) in a FAIR Metadata Registry for COVID-19 Research. Stud Health Technol Inform. 2021 Nov;287:73–7, doi: 10.3233/SHTI210817.

[7] Sansone SA, Rocca-Serra P, Field D, Maguire E, Taylor C, Hofmann O, et al. Toward interoperable bioscience data. Nat Genet. 2012 Feb;44(2):121–6, doi: 10.1038/ng.1054.

[8] HL7 International. HL7 FHIR Release 4 - Welcome to FHIR® [Internet]. 2019. Available from: https://www.hl7.org/fhir/R4/

[9] Wolstencroft K, Owen S, Krebs O, Nguyen Q, Stanford NJ, Golebiewski M, et al. SEEK: a systems biology data and model management platform. BMC Syst Biol. 2015 Jul;9:33, doi: 10.1186/s12918-015-0174-y

[10] NFDI4Health Task Force COVID-19, Darms J, Henke J, Hu X, Schmidt CO, Golebiewski M, et al. Improving the FAIRness of Health Studies in Germany: The German Central Health Study Hub COVID-19. In: Delgado J, Benis A, de Toledo P, Gallos P, Giacomini M, Martínez-García A, et al., editors. Studies in Health Technology and Informatics [Internet]. IOS Press; 2021 [cited 2022 Nov 30]. Available from: https://ebooks.iospress.nl/doi/10.3233/SHTI210818

[11] Sansone SA, Rocca-Serra P, Gonzalez-Beltran A, Johnson D, ISA Community. Isa Model And Serialization Specifications 1.0. 2016 Oct 28 [cited 2022 Nov 24]; Available from: https://zenodo.org/record/163640

[12] ISA-tools. ISAdatasets/json [Internet]. 2022. Available from: https://github.com/ISA-tools/ISAdatasets/tree/master/json

[13] JSON Schema. Specification. [Internet]. 2020 [cited 2023 Mar 06]. Available from: https:/json-schema.org/specification.html

[14] Klopfenstein SAI, Sass J, Vorisek CN, Jorczik F, Thun S, NFDI4Health Task Area 2. Implementation Guide - ISA to FHIR [Internet]. Available from: https://simplifier.net/guide/ISA-to-FHIR/Home?version=current

[15] Vorisek CN, Lehne M, Klopfenstein SAI, Mayer PJ, Bartschke A, Haese T, et al. Fast Healthcare Interoperability Resources (FHIR) for Interoperability in Health Research: Systematic Review. JMIR Med Inform. 2022 Jul 19;10(7):e35724. doi: 10.2196/35724.

[16] HL7 International. HL7 FHIR FHIR CI-Build - Welcome to FHIR® [Internet]. 2022. Available from: http://build.fhir.org/