

Analysis of ISO/TS 21526 Towards the Extension of a Standardized Query API

Hannes ULRICH^{a,1}, Ann-Kristin KOCK-SCHOPPENHAUER^a, Cora DRENKHAHN^a,
Matthias LÖBE^b and Josef INGENERF^{a,c}

^a *IT Center for Clinical Research (ITCR-L), University of Lübeck, Germany*

^b *Institute for Medical Informatics, Statistics and Epidemiology (IMISE), University of Leipzig, Germany*

^c *Institute of Medical Informatics, University of Lübeck, Germany*

Abstract. Metadata is often used for different tasks in the field of medical informatics: semantic description of data, quality validation, data integration, or information retrieval. Metadata definitions are captured and curated in time-consuming tasks and stored in metadata repositories that manage and preserve the metadata. Due to technical and legal restrictions, metadata is rarely as easily accessible and interoperable as it is necessary for modern information systems. In a previous study, a uniform interface based on the widely used ISO/IEC 11179 and the Facebook data retrieval language GraphQL was introduced as a solution to these technical obstacles. In the meantime, the ISO standard 21526 has been published, a recent version designed with a strong focus on health informatics. While it is conceptually oriented on the metamodel in ISO 11179, a number of extensions but also restructurings have been introduced. In this study, the authors investigated the difference between ISO 11179 and ISO 21526 and extended the unified metadata query interface to be future-proof and in particular, to support the semantic extensions of ISO 21526.

Keywords. Metadata, ISO 11179, ISO 21526, Metadata Repository, GraphQL

1. Introduction

Metadata – in our definition machine-readable descriptions of items of data - is increasingly applied in the field of medical informatics and is often used for different tasks, e.g. semantic characterization, quality validation, or data integration. Metadata definitions are captured and curated in time-consuming tasks, involving experts and data stewards to ensure reliability. The information is stored in metadata repositories (MDR) that manage and preserve the metadata. Due to technical and structural obstacles, metadata is rarely interoperable. This hampers aggregation and management of (meta-)data sets in order to answer research questions (1,2). One reason is, that the leading metadata standard ISO 11179 (3) does not constrain implementations, so existing interfaces of MDR systems differ technically. In earlier studies, Nguoungo et al. (4) and Park et al. (5) showed structural problems of ISO/IEC 11179. So, the recently published ISO 21526 standard, successor of ISO 11179 with focus on medical applications,

¹ Corresponding Author, Hannes Ulrich, IT Center for Clinical Research, Lübeck; Telephone: +49 (0) 451 - 3101 5607; Fax: +49 451 3101 5604; E-mail: h.ulrich@uni-luebeck.de

introduces new concepts, but also restructures existing concepts and aims to overcome the structural problems. This raises two new research questions: (1) does the restructuring create incompatibility between the two standards and (2) are the extended possibilities offered by the new standard profitable enough to integrate them into existing systems?

In this study, the authors investigate the difference between ISO 11179 and its successor ISO 21526. If the comparison shows remarkable enhancements, an extension of our standardized metadata interface will be present as part of this study to support the ISO 21526 and to be adaptable to upcoming systems.

2. Background

The ISO/IEC 11179 (3) is a much-used metadata norm in the field of medical informatics. The defined metamodel separates the representation of structural information from the conceptual categorization of metadata. The central information object, called data element, is defined by definitions and value domains that restrict the value represented, and by a link to data element concepts to describe its information in a semantic-preserving manner. Various MDR systems use the standard to constrain and harmonize their information: caDSR (6), METeOR (7), Aristotle (8), and USHIK (9). Since ISO 11179 does not constrain implementation, existing MDR systems differ technically in the provided interfaces. In a prior study, we designed QL⁴MDR (10) as a new approach to overcome technical heterogeneity of the existing MDRs. Inspired by HL7 FHIR and its uniform interface concept (11), QL⁴MDR and the underlying schema is an interface definition that constrains the exchanged data based on the ISO 11179-3 metadata model.

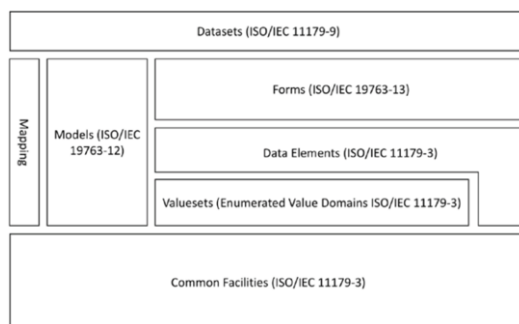


Figure 1. The metadata standards ISO 11179 and 19763 are combined in the new ISO 21526 to define metadata repository requirements. Additionally, a mapping package is introduced (12).

The successive ISO/TS 21526 Health informatics — Metadata repository requirements (MetaRep) is designed to be an extension and a clarification of 11179 (12). It combines the well-known metadata metamodel of 11179-3 with the metadata standard ISO 19763 and aims to simplify the definition of metadata despite its structural complexity, as shown in [Figure 1](#). It is focused on capturing the interrelation between data models, which are used to exchange information in healthcare. The storage of these interrelations and their contextual information are necessary for the later interpretation and (re)use of the exchanged data.

The authors will examine both standards systematically focusing on the research questions mentioned before. The analysis checks the packages of both standards against each other to identify effective modifications. Newly introduced packages and classes will be examined towards their changes regarding their impact on the limitations (4,5).

3. Results

3.1. Comparison of ISO 11179 and ISO 21526

In ISO 21526, various previous packages are aggregated to refocus on metadata definition resulting in size-wise reduction regarding numbers of defined packages and classes. The data description package of 11179-3 is still the core model of the new standard. But the successor introduced a third conceptual definition axis between Conceptual Domain and Value Domain to link external concept systems. ISO 21526 favors HL7 FHIR CodeSystems (13), including LOINC and SNOMED CT, to be used in new Conceptual Domain Definitions. The concept package is simplified class-wise and modeled according to the Simple Knowledge Organization System (SKOS) (14) to make implementation easier. As a novelty, ISO 21526 introduces a mapping model to provide a uniform way to describe mappings between (artifacts of) data elements. A mapping is defined as an association between two different items, characterized by a type with a value set of elements like *broader*, *narrow*, *related*, *same_as* and *derived_from*.

3.2. Expansion of QL^AMDR

The schema below was derived from ISO 21526 to match the newly introduced mapping classes and furthermore the classes of concept package were included to enable enhanced querying. As shown in **Figure 2**, the first schema contained six (plus seven supportive) objects and was expanded by five additional classes: Concept, SemanticRelation, Map, MDRMapping and Conceptual Domain Definition. Concept and MDRMapping are introduced as new entry points to start a query at these objects.

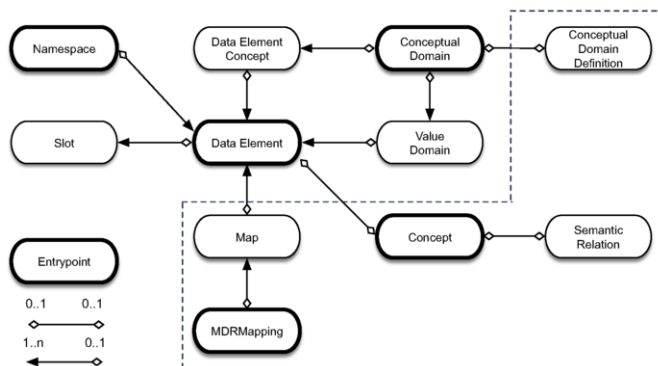


Figure 2. The newly added objects are shown in the dashed box. Besides the existing three entry points Conceptual Domain, Namespace, and Data Element, two new were added: MDRMapping and Concept.

4. Discussion

ISO 21526 is a constructive extension to the commonly used ISO 11179 but also inherits some problems. The third conceptual definition axis is a beneficial addition to the standard. It enables the direct usage of ontological knowledge as a conceptual domain as well as the usage of predefined and externally managed value sets for the value domain. The emerging adoption of FHIR and thereby provided machine-readable concept systems (in FHIR code systems and value sets) are beneficial for metadata repositories. In earlier studies, Ngouongo et al. (4) and Park et al. (5) described and categorized the problems of ISO/IEC 11179: the absence of semantic or syntactic linkage of shared concepts between components (15) and the missing support of structure for either metadata extension mechanism (2) or usage model (16). Remodeling the concept package towards SKOS and the introduction of the mapping package opens the possibility to solve the problem of missing linkage between concepts. The structural mapping using the MDRMapping and semantic annotation using concepts with SKOS enables direct links between every administrated item. The missing technical extensibility, like the FHIR extensions, is a structural problem and should have been addressed in the new metadata standard. Machine-readable or machine-actionable extensions are highly useful as demonstrated by the often-used FHIR extensions and profiles and recommended by the renowned FAIR principles (17). A structure for a usage model of the metadata is not directly addressed, but the newly added semantic possibilities allow SKOS-based annotations. Context and corresponding usage should not be annotated using domain-specific ontologies like SNOMED CT (18) since they are describing the “what is”, whereas the usage is dependent on the situation and its context. Conceptual orientation thus requires that each term in the vocabulary has a single, coherent meaning, even though its meaning may vary depending on its occurrence in a context (19). On the contrary, SKOS is able to represent this contextual relationship pragmatically and thus to depict a usage model.

The extension of the previously developed QL⁴MDR will enable data sharing between MDR of both ISO standards since the underlying metamodel is not altered. The introduction of the mapping class is beneficial for federated metadata processing. The standardized mapping between metadata items enables schema crosswalks between different items in different systems and promotes their reuse and sharing due to their findability. The upcoming implementations based on ISO 21526 will open up interesting possibilities, for example, for the consensus process of core data sets with preferred data elements based on multiple existing, possibly conflicting data set specifications in source systems. Additionally, an extension of QL⁴MDR does not break the current interface implementation due to the nature of GraphQL and can support semantic querying for metadata.

5. Conclusion

The new metadata standard ISO 21526 is a qualified successor and solves some inherited problems of the leading ISO 11179 using a good combination of newly introduced and refined packages. The extension of QL⁴MDR will enable better queries using semantic identifiers, and the mapping classes will be beneficial for the metadata processing, especially in a federated context.

Acknowledgements

This work was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft) DFG grants IN 50/3-2 and WI 1605/10-2.

References

- [1] Dugas M, Jöckel K-H, Friede T, Gefeller O, Kieser M, Marschollek M, et al. Memorandum “Open Metadata.” *Methods Inf Med.* 2015;54(4):376–8.
- [2] Nadkarni PM, Brandt CA. The Common Data Elements for cancer research: remarks on functions and structure. *Methods Inf Med.* 2006;45(6):594–601.
- [3] ISO/IEC. ISO/IEC 11179-3:2013 - Information technology -- Metadata registries (MDR) -- Part 3: Registry metamodel and basic attributes [Internet]. 2013. Available from: <https://www.iso.org/standard/50340.html>
- [4] Nguongo SMN, Löbe M, Stausberg J. The ISO/IEC 11179 norm for metadata registries: Does it cover healthcare standards in empirical research? *J Biomed Inform.* 2013 Apr;46(2):318–327.
- [5] Park YR, Kim JH. Achieving interoperability for metadata registries using comparative object modeling. In: *Studies in Health Technology and Informatics* [Internet]. 2010. p. 1136–1139. Available from: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-78649520866&doi=10.3233%2F978-1-60750-588-4-1136&partnerID=40&md5=f3da6f5ef8c6517837565b072c067fac>
- [6] Warzel DB, Andonyadis C, McCurry B, Chilukuri R, Ishmukhamedov S, Covitz P. Common Data Element (CDE) Management and Deployment in Clinical Trials. *AMIA Annu Symp Proc.* 2003;2003:1048.
- [7] Australian Institute of Health and Welfare. METeOR home [Internet]. [cited 2018 Jun 29]. Available from: <http://meteor.aihw.gov.au/content/index.phtml/itemId/181162>
- [8] Aristotle Metadata Registry - [cited 2019 Feb 20]. Available from: <https://www.aristotlemetadata.com/>
- [9] Penozo C, Barnes R, Chaparro M, Carney J, Levin B, Chudy M, et al. United States Health Information Knowledgebase (USHIK). *Int J Funct Inform Pers Med.* 2010;3(4):292–313.
- [10] Ulrich H, Kern J, Tas D, Kock-Schoppenhauer A-K, Ückert F, Ingenerf J, et al. QL4MDR: a GraphQL query language for ISO 11179-based metadata repositories. *BMC Med Inform Decis Mak.* 2019 Mar 18;19(1):45.
- [11] Bender D, Sartipi K. HL7 FHIR: An Agile and RESTful approach to healthcare information exchange. In: *Proceedings of the 26th IEEE international symposium on computer-based medical systems.* IEEE; 2013. p. 326–31.
- [12] ISO/IEC. ISO/TS 21526:2019 - Health informatics — Metadata repository requirements (MetaRep) [Internet]. ISO. 2019. Available from: <https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/07/10/71041.html>
- [13] Benson T, Grieve G. *Principles of Health Interoperability - SNOMED CT, HL7 and FHIR.* 3rd ed. London: Springer International Publishing; 2016.
- [14] Miles A, Bechhofer S. SKOS simple knowledge organization system reference. *W3C Recomm.* 2009;18:W3C.
- [15] Davies J, Harris S, Crichton C, Shukla A, Gibbons J. Metadata standards for semantic interoperability in electronic government. In: *Proceedings of the 2nd international conference on theory and practice of electronic governance.* 2008. p. 67–75.
- [16] Solbrig HR. Metadata and the reintegration of clinical information: ISO 11179. *MD Comput Comput Med Pract.* 2000;17(3):25.
- [17] Wilkinson et al. 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;3:160018.
- [18] Schulz S, Stegwee R, Chronaki C. Standards in healthcare data. In: *Fundamentals of Clinical Data Science.* Springer; 2019. p. 19–36.
- [19] Cimino J. Desiderata for Controlled Medical Vocabularies in the Twenty-First Century. *Methods Inf Med.* 1998 Nov;37(4–5):394–403.