German Medical Data Sciences 2023 — Science. Close to People.
R. Röhrig et al. (Eds.)

doi:10.3233/SHTI230721

Mettertron – Bridging Metadata Repositories and Terminology Servers

Jan SCHLADETZKY [a,1], Ann-Kristin KOCK-SCHOPPENHAUER [a], Cora DRENKHAHN [a,b], Josef INGENERF [a,b], and Joshua WIEDEKOPF [a]
[a] *IT Center for Clinical Research, University of Lübeck, Lübeck, Germany*
[b] *Institute for Medical Informatics, University of Lübeck, Lübeck, Germany*

Abstract. To provide clinical data in distributed research architectures, a fundamental challenge involves defining and distributing suitable metadata within Metadata Repositories. Especially for structured data, data elements need to be bound against suitable terminologies; otherwise, other systems will only be able to interpret the data with complex and error-prone manual involvement. As current Metadata Repository implementations lack support for querying externally defined terminologies in FHIR terminology servers, we propose an intermediate solution that uses appropriate annotations on metadata elements to allow run-time Terminology Services mediated queries of that metadata. This allows a very clear separation of concerns between the two related systems, greatly simplifying terminological maintenance. The system performed well in a prototypical deployment.

Keywords. Interoperability; Metadata Repositories; Terminology Servers

1. Introduction

The ongoing digitalization of the healthcare sector is leading to rapid growth the amount and depth of clinical data. The secondary use of this data plays a crucial role in enhancing research and improving healthcare costs [1] but places substantial demands on data interoperability [2,3]. Syntactic and semantic interoperability is fundamental in achieving pervasive digitalization in medicine and enabling cross-institutional collaborations. For the semantic level, metadata repositories (MDRs) and terminology servers (TS) are essential tools. MDRs manage the metadata of data elements, while TSs manage clinical terminologies, classifications, coding schemes, and use-case-specific value sets.

However, historically, MDRs and TS haven't been considered related systems and have undergone separate development and research cycles. Hence, it is currently required to maintain terminology in the MDR itself, alongside the leading TS terminology, to use the tools an MDR offers for data integration. This results in significant work for data stewards and depends on update workflows to maintain consistency and synchronicity between the systems.

Hence, in a previous work, we have already examined a similar use case, where we enabled a FHIR TS to manage the terminology available in the MDR. This system, called

[1] Corresponding Author: Jan Schladetzky, Universität zu Lübeck, Ratzeburger Allee 160, 23562 Lübeck, Germany; Telephone: +49 451 500-52702; E-Mail: j.schladetzky@uni-luebeck.de.

TermiCron [4], delegates the task of terminology maintenance to the FHIR TS by creating the required data structures by the mechanisms provided by the MDR. This allows the use of FHIR ValueSets stored on a terminology server as catalogs in the MDR implementations and thus enables needed API functionality.

The present work significantly expands on this idea in fully delegating terminological tasks, no longer requiring the provision of terminology in the MDR, and thus enforcing a clear split of responsibilities of the two related but distinct systems. This approach aims to develop a system for linking MDRs to TSs, to achieve a separation of concerns, provide better support for complex terminologies, and consequently improve data quality.

2. Methods

Terminology Binding

A fundamental challenge in the modeling process of clinical data is the binding of data elements to suitable terminologies [5,6]. This necessitates the creation and maintenance of use-case specific ValueSets (VS), which select codes from suitable CodeSystems (CS) or other VS. Data elements are then bound to VS. The FHIR terminology module has become the leading mechanism in accomplishing this task [2]. Where possible, modelers use existing clinical terminologies, such as the ICD family, SNOMED CT, RadLex, and others [2,7]. Specific use cases often require mappings between terminological systems maintained in ConceptMap (CM) resources, especially for local, non-standard codes mapped to standard terminology. The design, maintenance, and distribution of these terminological resources impedes constructing a healthcare IT infrastructure [3,7].

Since the description of a dataset in an MDR should also maintain the terminology attached to these data elements, the corresponding VSs need to be provided by the MDR systems. The naïve approach towards providing that data requires creating MDR catalogs from the VS. However, this requires a significant maintenance effort to manually track changes to the VS, which are often defined and updated in external processes and maintained in (FHIR-based) terminology servers [8,9].

Our previous work, *TermiCron* [4], represents an important first step towards a clean separation of concerns but provides additional challenges that arise when the FHIR VS definition changes.

ISO/TS 21526

There are two primary ISO standards relevant to MDRs: ISO/IEC 11179 [10] and ISO/TS 21526 [11]. ISO/IEC 11179 standardizes the representation of metadata in a metadata registry while ISO/TS 21526 focusses on the healthcare sector and specifies extensions to ISO/IEC 11179 and other standards. It also suggests simplifications and restrictions to ISO/IEC 11179 for the healthcare sector. It is important to note that ISO/TS 21526 directly anticipates the binding of data elements against FHIR VS and the external maintenance of these VS in a separate FHIR TS [11,12].

3. Results

We have implemented a middleware to allow TS-driven terminology binding in MDR implementations that don't natively offer this functionality. *Mettertron* acts as a proxy between a requesting client and an MDR in conjunction with at least one TS. *Mettertron* currently offers two REST API endpoints, one for the *$validate-code* and one for the *$translate* operation as defined in the FHIR standard, both illustrated in Figure 1.

Through these endpoints, *Mettertron* enables requests to an MDR to use the services a terminology server provides. *Mettertron* processes HTTP requests to the API in a three-step process: When receiving a client request, we query the MDR for the attributes of the specified data element, which needs to be uniquely identified in the client's request. These attributes identify resources in the TS that are relevant to the data. In the second step, a request for the respective FHIR operation is sent to the TS with the information gathered from the element attributes. Finally, the response by the TS is parsed and sent back to the initial requester. Downstream requests are cached for increased performance.

MDR elements' attributes

We defined several attributes for representing the terminological parameters of data elements in an MDR. With the help of these attributes, it is possible to specify the

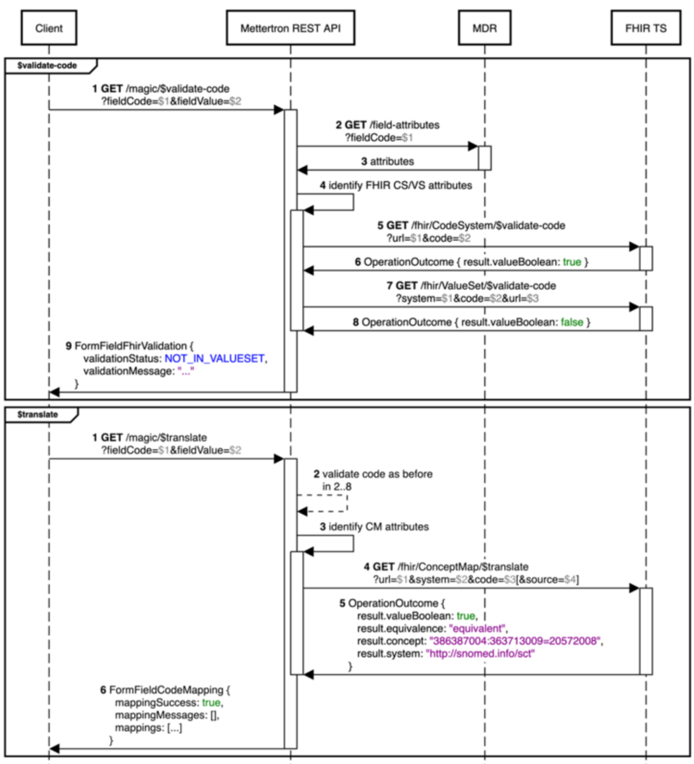

Figure 1: Sequence diagram showing the *$validate-code* and *$translate* endpoints of Mettertron. The system requires only the parameters for the identification of the data element and the instance data, and Mettertron automatically calls the MDR as required.

canonical URLs of the value sets, code systems, and concept maps for linking to external terminology. Since it is possible in all MDRs to define custom attributes, the names of the attributes used by *Mettertron* are configurable as well.

Code validation

One important FHIR operation is the *$validate-code* operation, allowing the verification that a code belongs to a specified VS or CS. We require that the attribute for the CS level is set and will always validate on that level, while VS validation is only carried out if the respective attribute is present.

Automatic code translation

The FHIR operation *$translate* allows the translation of codes from a source code system to one or more target code systems using the defined relationships by the CM resource. The operation requires CS, VS, and CM attributes in the MDR to a) state the context of the mapping, b) provide the source and target CS URL, and c) to reference the actual CM resource that determines the mapping.

Adaptability

Mettertron is written in Kotlin using the *Ktor* framework and thus runs on the Java JVM. It was developed to support different API structures of other MDR and TS implementations and reduce the amount of work needed to adapt new systems by offering generic interfaces. The application's source code is available under the terms of the GNU Affero General Public License via GitHub (github.com/itcr-uni-luebeck/mettertron) and Zenodo (10.5281/zenodo.7373740).

4. Discussion

We have performed two deployments using the Ontoserver FHIR TS [8] and two different MDRs: the commercially available CentraXX MDR [13] and the open-source Data Element Hub [14]. While the FHIR TS is interchangeable, due to different API structures, we currently don't offer support for other MDRs. This is the subject of ongoing development. Our prototypes operated well in terms of performance and validity in a scenario with a data element bound against ICD-10-GM and a corresponding VS; the same applies to a scenario using the more complex reference terminology SNOMED CT. The performance of terminological queries is a function of the chosen TS since not all TS support the more complex mechanisms offered by SNOMED CT and other CS with the same performance.

MDRs play an essential role by documenting the requirements of source and target systems, thus enhancing semantic interoperability. This metadata must always be up to date. However, (terminology) maintenance is a time-consuming task, especially in the case of terminologies with complex structures. Such resources essentially require the use of a terminology server. Our study aims at bridging the gap between these two systems.

In our opinion, external systems like *Mettertron* and *TermiCron* shouldn't be required for this use case, and we anticipate the development of entirely ISO 21526-compliant systems that offer this vital functionality natively.

We have observed a hesitant adoption of MDRs in research data projects like the Medical Informatics Initiative in Germany [15], with few partners choosing to integrate MDRs into their Data Integration Centers. The reason for this may lie partly with the amount of work needed not only in representing data structures in the MDR but also in manually providing the bound terminologies to that system.

5. Conclusions

Mettertron facilitates the combined use of MDR and TS through its API endpoints. This can enable the use of a TS in cases where the used MDR does not offer satisfying support for terminological services but is clearly a bridge technology until MDRs provide support for this functionality.

Nevertheless, a solution like the one discussed in this paper represents an important tool for simplifying the deployment and maintenance of MDRs. Our approach to bridging the gap between MDRs and TSs enables the sharing and reuse of metadata a. It thus helps achieving the overarching goal of interoperability in the healthcare sector.

Acknowledgments

This work was funded by the German Federal Ministry of Education and Research (BMBF) as part of the Medical Informatics Initiative Germany, grant 01ZZ2312A.

Conflict of Interest

The authors declare that there is no conflict of interest.

Contributions of the authors

JW wrote the first version of the software, JS implemented support for the DEHub. JS and JW wrote the first version of the manuscript. JS, JW, AK and CD revised the manuscript. JI supervised the research. All authors have approved the manuscript as submitted and accept responsibility for the scientific integrity of the work.

References

[1] Boyd M, Zimeta DM, Tennison DJ, Alassow M. Secondary use of health data in Europe. Open Data Institute; 2021.
[2] Ingenerf J, Drenkhahn C. Referenzterminologie SNOMED CT – Interlingua zur Gewährleistung semantischer Interoperabilität in der Medizin. 1st ed. Springer Vieweg Wiesbaden; 2023.
[3] Lehne M, Sass J, Essenwanger A, Schepers J, Thun S. Why digital medicine depends on interoperability. Npj Digit Med 2019;2:1–5.
[4] Wiedekopf J, Ulrich H, Drenkhahn C, Kock-Schoppenhauer A-K, Ingenerf J. TermiCron - Bridging the Gap Between FHIR Terminology Servers and Metadata Repositories. Stud Health Technol Inform 2022;290:71–5.
[5] Rector AL, Qamar R, Marley T. Binding ontologies and coding systems to electronic health records and messages. Appl Ontol 2009;4:51–69.
[6] Rossander A, Lindsköld L, Ranerup A, Karlsson D. A State-of-the Art Review of SNOMED CT Terminology Binding and Recommendations for Practice and Research. Methods Inf Med 2021;60:e76–88.

[7] Benson T, Grieve G. Principles of Health Interoperability: FHIR, HL7 and SNOMED CT. 4th ed. Springer International Publishing; 2021.
[8] Metke-Jimenez A, Steel J, Hansen D, Lawley M. Ontoserver: a syndicated terminology server. J Biomed Semant 2018;9:24.
[9] Cimino JJ. Desiderata for Controlled Medical Vocabularies in the Twenty-First Century. Methods Inf Med 1998;37:394–403.
[10] ISO/IEC. ISO/IEC 11179-3:2023 - Information technology -- Metadata registries (MDR) -- Part 3: Registry metamodel and basic attributes 2023.
[11] ISO/IEC. ISO/TS 21526:2019 - Health informatics — Metadata repository requirements (MetaRep). ISO 2019.
[12] Ulrich H, Kock-Schoppenhauer A-K, Deppenwiese N, Gött R, Kern J, Lablans M, et al. Understanding the Nature of Metadata: Systematic Review. J Med Internet Res 2022;24:e25440.
[13] KAIROS GmbH. CentraXX MDR 2022. https://www.kairos.de/produkte/centraxx-mdr/ (accessed November 28, 2022).
[14] Kadioglu D, Vengadeswaran A, Folz M, Barakat A. DataElementHub (DEHub) n.d.
[15] Semler SC, Wissing F, Heyder R. German Medical Informatics Initiative. Methods Inf Med 2018;57:e50–6.