A Methodology for an Auto-Generated and Auto-Maintained HL7 FHIR OWL Ontology for Health Data Management

Vassilis KILINTZIS[a,1], Vasileios C. ALEXANDROPOULOS[a],
Nikolaos BEREDIMAS[a] and Nicos MAGLAVERAS[a]

[a] *Laboratory of Computing, Medical Informatics and Biomedical Imaging Technologies, Medical School, Aristotle University of Thessaloniki, Greece*

Abstract. The process of maintenance of an underlying semantic model that supports data management and addresses the interoperability challenges in the domain of telemedicine and integrated care is not a trivial task when performed manually. We present a methodology that leverages the provided serializations of the Health Level Seven International (HL7) Fast Health Interoperability Resources (FHIR) specification to generate a fully functional OWL ontology along with the semantic provisions for maintaining functionality upon future changes of the standard. The developed software makes a complete conversion of the HL7 FHIR Resources along with their properties and their semantics and restrictions. It covers all FHIR data types (primitive and complex) along with all defined resource types. It can operate to build an ontology from scratch or to update an existing ontology, providing the semantics that are needed, to preserve information described using previous versions of the standard. All the results based on the latest version of HL7 FHIR as a Web Ontology Language (OWL-DL) ontology are publicly available for reuse and extension.

Keywords. HL7 FHIR, OWL, RDF

1. Introduction

As population health management (PHM) becomes the new best practice of healthcare, a new information technology infrastructure is needed to facilitate this care delivery model. Within this infrastructure electronic health records (EHRs) are necessary but not sufficient. Interoperability among health IT systems and with other data sources is crucial to PHM but is still far from being achieved. Interoperability advocates are promoting the development of the Fast Health Interoperability Resources (FHIR) standard and the use of open application programming interfaces (APIs) [1].

Semantic technologies and Linked Data principles [2] are in the heart of the solution that entails multiple sources of data and information along with multiple access points, a strong temporal aspect, as well as different computational workflows.

FAIR Principles emphasize on the capacity of computational systems to find, access, interoperate, and reuse data with none or minimal human intervention [3]. To

---

[1] Corresponding Author, Vassilis Kilintzis; E-mail: billyk@med.auth.gr.

that end, the use of standards and definition of detailed semantics, as early as possible, in the process of domain definition is critical.

Several public or commercial repositories are modeling knowledge using semantic web technologies. One of the easiest to use and freely accessible repository of health-related ontologies is Bioportal of the National Center for Biomedical Ontology [4]. It incorporates search and representation mechanisms for several health ontologies and terminologies such as SNOMED Clinical Terms, International Classification of Diseases (ICD) and Logical Observation Identifier Names and Codes (LOINC) among others. There is additional work being done in the process of transforming OWL ontologies into FHIR terminology resources as presented in [5]. The paper presents the challenges in the transformation a detailed overview of the mapping between OWL and FHIR.

In our previous work we have presented, initially, a representation as an ontology of the HL7 FHIR primitive and complex data types [6], then, a platform to support integrated care built upon linked data principles based on an ontology representation of HL7 FHIR[7] along with a methodology of maintenance of this ontology, aiming to keep up with the evolving standard and at the same time retain backwards compatibility with the software built upon previous versions[8].

In this paper we present the methodology used to develop a FHIR ontology generator software used for automatically transforming the FHIR specification resource files into an OWL ontology. The resulting ontology is ready to support the data management of EHR and PHM data along with the provision for managing future evolution of the standard. It includes all the FHIR defined data types and resources and, can be further expanded or restricted using owl axioms to adhere to a specific domain concepts or restriction on data. The ontology is shared publicly and aims on reusability by providing access to a very broad and componentized data model, helping researchers and adopters to overcome the triviality of re-implementing the base health-domain model each time.

2. Methods

The FHIR standard defines three types of data structures. Primitive data types, which are simple values, like string or decimal, complex data types, which are reusable collections of primitive types or other complex types, like *ContactDetails* or *HumanName*, and resources, which represent specific parts of the medical process and contain the data required to define them, like *Observation* or *AllergyIntolerance*.

In our ontology, each of the three data structure types are defined as an *owl:Class*. While primitive types could potentially be defined as *owl:Datatype*, *owl:Class* is used due to the absence of software support for custom data types. The properties of complex types and resources are mapped to *owl:ObjectProperty* in our ontology. Further details on the design decisions of the ontology can be found in [6] and [8].

2.1. FHIR Ontology Generator

FHIR provides various metamodels of the data structures defined as part of it[2]. For our needs, we use a JSON document defining the complex types and resources as instances

---

[2] https://www.hl7.org/fhir/downloads.htm

of *StructureDefinition* which is a meta-resource defined by FHIR to facilitate the exchange of custom resources. Primitive types are for the most part are "set in stone", so they are not created as a part of the auto generator. An existing implementation is instead added to the created ontology. The simple mappings are presented in Table 1.

Property "kind" can have one of four values "logical", "resource", "complex-type", "primitive-type". Values "logical" and "resource" have similar semantics and are treated the same. Value "primitive-type" is ignored as discussed previously.

Property "name", is used as the URI suffix of the resulting *owl:Class*.

The last property "differential.element" of *StructureDefinition*, is an array of *ElementDefinition*. *ElementDefinition* is a meta-resource that defines a single element of the modelled resource (e.g. the semantics of "*Observation.referenceRange.type*" element of the *"Observation"* resource are described as an *ElementDefinition*). Each *ElementDefinition* in the array is mapped to an *owl:ObjectProperty* assigned via *rdfs:domain* to the specific *owl:Class*. The semantics accompanying the specific *ElementDefinition* are defined at the corresponding *owl:Class*.

The property "path" is used to identify the *rdfs:domain* of the *owl:ObjectProperty*. In FHIR, resources often contain elements that need to be grouped together in a single substructure. When a substructure, like this, is defined only in the context of a single resource, it is modeled as an extension of the complex type *BackboneElement*. Depending on the value of "path" the *rdfs:domain* is either an existing *owl:Class* (corresponding to either a resource or of a complex-type), or a new *owl:Class* (*rdfs:subClassOf FHIRct:BackboneElement*). Property "path" is also used to generate the *rdfs:label* of the *owl:ObjectProperty*.

Property "type" includes one or many substructures that determine the allowed values for this element. There are two cases. In one case, when the allowed values can be a simple or complex type, the value becomes the *rdfs:range* of the property. In the other case the allowed value is a reference to a resource so the allowed types become the *rdfs:range* and the custom annotation *takesReference* is added to the *owl:ObjectProperty* to allow easier handling in applications using it.

Table 1. FHIR to OWL mapping

FHIR StructureDefinition	FHIR ElementDefinition	OWL Namespace:Local Name
name	path	rdfs:label
version		owl:version
description	definition	rdfs:comment
baseDefinition		rdfs:subClassOf
differential.element		owl:ObjectProperty
	min, max	owl:minCardinality, owl:maxCardinality, owl:cardinality
	isSummary	pt:isSummary
	isModifier	pt:isModifier

pt is the namespace: *http://lomi.med.auth.gr/ontologies/FHIRPrimitiveTypes#*

2.2. Integrating version changes to existing ontologies

The ontology generator described in 2.1 is sufficient when building an ontology based on the latest version of FHIR from the ground up. When a new version of the standard is introduced and since ontologies of previous versions could be already used, updates in the resources or complex data types can cause problems. A methodology for managing these possible problems, aiming to maintain existing data and their semantics, even if they were defined by previous versions of the standard, using the

expressivity provided by OWL is presented in [8]. The presented generator is using a FHIR-provided JSON document with the changes between versions to tackle this issue automatically. The useful types of changes documented are the following:

- Addition or deletion of elements, or name change to elements
- Changes to minimum and maximum cardinality of a property
- Changes to allowed types of a property

The JSON document contains an object for each FHIR data structure with a property "status". In the case of "changed", an object "elements" is provided that has a key for each of the changed properties of this particular data structure. The value of each key in "elements" is an object that may contain one or more definitions. The possible combinations are shown in Table 2.

Table 2. Handling of changes between versions

Status	Property combination	Changes to owl:Class	Changes to owl:ObjectProperty
deleted		Add owl:deprecated	Add owl:deprecated
changed	old-min – new-min	Change owl:minCardinality	
	old-max – new-max	Change owl:maxCardinality	
	removed-types		Remove from rdfs:range
	added-types		Add to rdfs:range
new		Create new Class	Create new ObjectProperty

3. Results

The generator is implemented as a Java app to take advantage of the Apache Jena API for building the ontology. Upon execution the corresponding ontology is created as an RDF/turtle file with the following base URIs:

http://lomi.med.auth.gr/ontologies/FHIRComplexTypes
http://lomi.med.auth.gr/ontologies/FHIRPrimitiveTypes
http://lomi.med.auth.gr/ontologies/FHIRResources

The procedure of the ontology update was tested using FHIR Release 3 (STU) and the current version FHIR Release #4.

The OWL-DL ontology automatically produced by the generator from FHIR v4.0.1: R4, is available for review and reuse in http://lomi.med.auth.gr/ontologies/FHIR

4. Discussion

While a semantic data model represented as an OWL ontology is an obvious choice to provide semantics to the modeled entities, and efforts of transforming relational databases to ontologies and vice-versa are widespread [9], using an ontology to describe, store and exchange actual health related data is not as common. Having HL7 FHIR as the de facto standard for exchanging health data, a system based on an ontology, derived from FHIR, could minimize data mappings and transformations, semantically enrich the managed health data to tackle interoperability obstacles and enhance findability, as proposed by FAIR principles. Such a system presented in [8]

requires maintenance of the FHIR based ontology and this process requires a lot of effort and attention.

In this paper we have presented the FHIR ontology Generator software that is capable of automatically transforming the provided by the standard definition files to an OWL-DL ontology. The resulting ontology includes all the semantics needed to support EHR data storage with data validation (i.e., accepted values, cardinality of values) along with support for automatic application into the ontology of future changes in the standard, maintaining intra-model semantics for backwards compatibility.

FHIR ontology Generator software can reduce the interoperability gap exhibited among health record systems and offer a model for PHM by providing either a solid data model base for specific domain implementations or a reference model to map existing deployed models. In the first option, the data model base, generated from the latest version of HL7 FHIR standard, can be restricted to accept specific domain entities with additional semantics, while in the second options, mapping to the model can be used, as a first step in the pipeline, to export FHIR resources from non-FHIR based EHR system.

Next steps include the integration, in the software, of the transformation rules that aim at the in-FHIR defined terminologies and bindings, as well as the development of a new software for the export of ontology instances as JSON FHIR resources.

Acknowledgements

References

[1]   Watson Health, Population health management beyond the EHR: Part 2 - Watson Health Perspectives, *IBM Watson Health Perspectives*, 2017. https://www.ibm.com/blogs/watson-health/population-health-management-beyond-ehr-part-2/ (accessed Apr. 21, 2021).
[2]   Berners-Lee T, Linked Data, 2011. https://www.w3.org/DesignIssues/LinkedData.html (accessed Feb. 18, 2015).
[3]   FAIR Principles, https://www.go-fair.org/fair-principles/ (accessed Jul. 1, 2021)
[4]   Whetzel PL, et al. BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications. Nucleic acids research. 2011 Jun 14;39(suppl_2):W541-5. .
[5]   Metke-Jimenez A, Lawley M, Hansen D. FHIR OWL: Transforming OWL ontologies into FHIR terminology resources. InAMIA Annual Symposium Proceedings 2019 (Vol. 2019, p. 664). American Medical Informatics Association..
[6]   Beredimas N, Kilintzis V, Chouvarda I, Maglaveras N. A reusable ontology for primitive and complex HL7 FHIR data types. In2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2015 Aug 25 (pp. 2547-2550). IEEE.
[7]   Kilintzis V, Chouvarda I, Beredimas N, Natsiavas P, Maglaveras N. Supporting integrated care with a flexible data management framework built upon Linked Data, HL7 FHIR and ontologies. Journal of biomedical informatics. 2019 Jun 1;94:103179.
[8]   Kilintzis V, Kosvyra A, Beredimas N, Natsiavas P, Maglaveras N, Chouvarda I. A sustainable HL7 FHIR based ontology for PHR data. In2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2019 Jul 23 (pp. 5700-5703). IEEE.
[9]   Spanos DE, Stavrou P, Mitrou N. Bringing relational databases into the semantic web: A survey. Semantic Web. 2012 Jan 1;3(2):169-209.