Building Standardized Semantic Web RESTful Services to Support ICD-11 Revision

Guoqian Jiang¹, Harold R. Solbrig¹, Christopher G. Chute¹

¹ Department of Health Sciences Research, Division of Biomedical Statistics & Informatics, Mayo Clinic College of Medicine, Rochester, MN. U.S.A.

{jiang.guoqian, solbrig.harold, chute}@mayo.edu

Abstract. For the beta phase of the 11th revision of International Classification of Diseases (ICD-11), the World Health Organization (WHO) intends to accept public input through a distributed model of authoring. To enable social computation and collaboration among broader communities, it is imperative but remains challenging to have an open-data access solution for ICD-11 contents. The objective of the present study is to describe our experiences and challenges on building the Simple Knowledge Organization System (SKOS) compliant Semantic Web RESTful services that aim to support the authoring application development of ICD-11. We propose a framework of building such services, which is composed of four layers: a semantic normalization layer, a semantic repository layer, a semantic services layer and an application layer. For the prototype implementation, we first generated the normalized ICD-11 linked RDF data through utilizing WHO ICD URI scheme and SKOS RDF data model. We used the Linked Data API and developed simple RESTful services for the standardized ICD-11 contents. We performed a preliminary evaluation on the usefulness of the system in terms of the feasibility of the services and its potential in support of building the Common Terminology Services 2 (CTS2) compliant services. In conclusion, we developed an effective framework of building standardized Semantic Web RESTful services on ICD-11 linked data, which would be useful to support ICD-11 revision.

Keywords: ICD-11 Revision, Common Terminology Services 2 (CTS2), Linked Open Data (LOD), Simple Knowledge Organization System (SKOS).

1 Introduction

The 11th revision of International Classification of Diseases (ICD-11) was officially launched by the World Health Organization (WHO) in March 2007 [1]. Historically, ICD was developed to support international comparison of mortality statistics. WHO has embraced a broadened set of use cases to drive ICD-11 development, including scientific consensus of clinical phenotype (definition and criteria), public health surveillance (e.g. mortality and morbidity) and clinical data aggregation [2]. Instead of using a traditional manual revision process leveraging international expertise, the ICD-11 revision aims to digitize the content of ICD in a scientific way with modern knowledge representation standards.

The beta phase of the ICD-11 revision started in May 2012, and WHO intends to accept public input through a distributed model of authoring. For the beta phase, there are a number of major challenges to the existing authoring framework and workflow [3-4]. First, the user community will potentially be large. Users will want to be able to access and download target ICD-11 contents into their own evaluation tools and/or try alternatives in their own environment. Second, the system must enable a sophisticated user role definition with a hierarchy of authority and a fine-grained permission matrix. This will be very important for supporting efficient social computing and for building communities in collaboration. Third, the system must support an open-data solution to lower the barriers for ICD data access. This openness will form the basis to enable human social computation and will empower the users to create their own views of data to satisfy different job roles and changing needs [5]. The present study is focused on the third challenge, i.e. building an open-data solution for ICD-11 data access.

Semantic Web technology provides a scalable framework for facilitating semantic data integration of heterogeneous resources and enabling semantic sharing through the standard query services. In particular, an emerging Semantic Web technology, known as Linked Open Data (LOD), has the potential to enable a revolution in how data is accessed and utilized. In the biomedical domains, Linked Data principles and practices have been adopted by an increasing number of projects. For instance, Linking Open Drug Data (LODD) is a task force within the W3C Health Care and Life Sciences Interest Group (HCLS IG). LODD has created Linked Data representations of the data sets about drugs, and identified interesting scientific and business questions that can be answered once the data sets are connected [6]. The National Center of Biomedical Ontology (NCBO) is embracing Semantic Web technologies as a common backend, and to extend the scope of its ontology access services to include RDF/OWL capabilities [2, 7]. Scalable and standardized knowledge bases are beginning to come into play, as typified by the repository of adverse drug events (ADEs) using the Linked Data technology [8]. In this study, we consider that the Linked Data technology will provide a scalable framework and solution for dealing with the challenge on the open data access of ICD-11 content.

The Simple Knowledge Organization System (SKOS), a W3C recommendation and standard, is a common data model for sharing and linking knowledge organization systems via the Web [9]. The SKOS data model provides a standard, low-cost migration path for porting existing knowledge organization systems to the Semantic Web.

The objective of the present study is to describe our experiences and challenges on building SKOS compliant Semantic Web RESTful services that aim to support the authoring application development of ICD-11. We propose a framework of building such services, which is composed of four layers: a semantic normalization layer, a semantic repository layer, a semantic services layer and an application layer. For the prototype implementation, we first generated the normalized ICD-11 linked RDF data through utilizing WHO ICD URI scheme and SKOS RDF data model. We used the Linked Data API and developed simple RESTful services for the standardized ICD-11 contents. We performed a preliminary evaluation on the usefulness of the system in terms of the feasibility of the services and its potential in support of building the Common Terminology Services 2 (CTS2) compliant services.

2 Background

2.1 ICD-11 Content Model

The purpose of ICD-11 content model is to present the knowledge that underlies the definitions of an ICD entity. The content model is composed of three layers: a foundation component, a linearization component and an ontological component [10]. The foundation component stores the full range of knowledge of all classification units in ICD. The linearization component corresponds to the classical print versions of ICD. The ontological component provides references to formal definition of terms and relationships. For the ontological component, WHO has sought to reuse existing ontologies including the Systematized Nomenclature of Medicine - Clinical Terms (SNOMED CT). SNOMED CT is the most comprehensive clinically oriented medical terminology system. It is owned and maintained by the International Health Terminology Standard Development Organization (IHTSDO). The IHTSDO and the WHO signed a collaborative agreement in July 2010, aiming at enabling harmonization - as complementary tools - of WHO Classifications and SNOMED CT, which essentially establishes SNOMED CT as the core of the ontological component of ICD [11]. The distributed authoring process envisioned for ICD will focus on the foundation layer, drawing on SNOMED terms for definitional context and populating the details of a rubric such as anatomy, severity, histopathology, or symptoms. Attributes within the foundation layer will enable the algorithmic derivation of a linearization from the foundation layer, thus unifying all linearizations with a common foundation core [2].

2.2 ICD URI Scheme

A Uniform Resource Identifier (URI) is commonly used to name and identify a resource, which has been treated as one of core principles of Semantic Web Linked Open Data by Tim Berners-Lee [12]. A Restful, resource-oriented service exposes a URI for every piece of data the client might want to operate on. A well-designed URI structure/scheme allows users easily explore and extend ICD data through an implementation of the RESTful services. Recently, an ICD URI scheme is proposed for naming and supporting web services by WHO. A base URI of http://id.who.int/icd/schema as the prefix for the vocabulary terms that related to ICD classification efforts maintained by WHO, http://id.who.int/icd/schema as the prefix for the vocabulary terms that related to ICD classification efforts maintained by WHO, http://id.who.int/icd/schema as the prefix for the vocabulary terms that related to ICD classification efforts maintained by WHO, http://id.who.int/icd/schema as the prefix for the vocabulary terms that related to ICD classification efforts maintained by WHO, http://id.who.int/icd/schema as the prefix for the vocabulary terms that related to ICD classification efforts maintained by WHO, http://id.who.int/icd/schema as the prefix for the vocabulary terms that related to ICD classification efforts maintained by WHO, http://id.who.int/icd/schema as the prefix for the vocabulary terms that related to ICD classification efforts maintained by WHO, http://id.who.int/icd/schema as the prefix for the vocabulary terms that related to ICD classification efforts maintained by WHO.

2.3 SKOS RDF Model

The SKOS data model views a knowledge organization system as a concept scheme comprising a set of concepts [9]. The vocabulary used in the SKOS data model is a set of URIs that specifies the notion of SKOS concepts, concept schemes, lexical labels, notations, documentation properties and semantic relations. SKOS data are expressed as RDF triples. An increasing number of SKOS datasets in RDF are publicly available. A typical example is the SKOS version of AGROVOC Agricultural Thesaurus provided by the Food and Agriculture Organization (FAO) of the United Nations (http://aims.fao.org/).

2.4 CTS2 Standard

The Common Terminology Services 2 (CTS2) standard defines the functional requirements of a set of service interfaces to allow the representation, access, and maintenance of terminology content either locally, or across a federation of terminology service nodes [13]. Note that the CTS2 standard has also incorporated the principles of Representational State Transfer (REST) web service in its specification. The CTS2 platform independent model (PIM) is rendered in the Unified Modeling Language (UML), which is publicly available for download [9]. The CTS2 PIM model describes the formal model and semantics of the various components. In the present study, we evaluate the potential of the SKOS model in support of building CTS2 compliant services on ICD-11 linked data.

2.5 Linked Data API

Linked Data API is a configurable API specification intended to support the creation of simple RESTful APIs over RDF triple stores [14]. The API specification is designed to be deployed as a proxy in front of a SPARQL endpoint to support: 1) generation of documents (or information resources) for the publishing of Linked Data; 2) provision of sophisticated querying and data extraction features, without the need for end-users to write SPARQL queries; 3) delivery of multiple output formats from these APIs, including a simple serialization of RDF in JSON syntax. Figure 1 shows system architecture of the Linked Data API. A prototype implementation of the Linked Data API, known as Elda, is publicly available at [15].

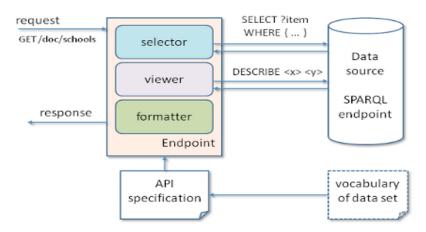


Figure 1. System architecture of the Linked Data API (Source from : <u>http://code.google.com/p/linked-data-api/</u>)

3 System Architecture

Figure 2 shows the system architecture of our proposed approach for building a standardized Semantic Web services for ICD-11 revision. The system is composed of the following four layers.

- Semantic Normalization Layer. This layer will provide normalization services against ad hoc OWL/RDF rendering of ICD-11 contents. The normalization should at least include: generating uniform IDs for ICD-11 categories using global identification model proposed by WHO; normalizing lexical properties of ICD-11 contents using the SKOS RDF model.
- 2) Semantic Repository Layer. In this layer, we will implement a RDF triplestore and establish a SPARQL server that provides a SPARQL endpoint. Note that it would be possible to expose the SPARQL endpoint externally for certain Semantic Web applications. This depends on the needs of the community.
- 3) Semantic Services Layer. In this layer, we will build light-weight Semantic Web RESTful services through the Linked Data API. The services will deliver multiple output formats, including traditional JSON and XML formats and the Semantic Web RDF and Turtle formats. Integrating the CTS2 model will produce the CTS2 compliant services, adding a layer of standardization for advanced applications.
- 4) Application Layer. ICD-11 authoring applications will utilize the Semantic Web RESTful Services to access ICD-11 contents. The existing authoring applications include iCAT alpha tool, WHO alpha browser, WHO beta browser, Exhibit browser, and Mayo's Proposal generator. The requirements solicited from the applications will iteratively define and improve Web services as described in the Services Layer above.

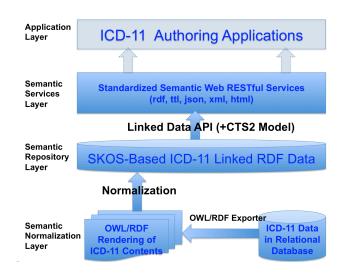


Figure 2. System architecture of our proposed approach for building a standardized Semantic Web services for ICD-11 revision.

4 Prototype Implementation

In the prototype implementation, we focused on normalizing the ICD-11 linked data using the SKOS-based RDF model and building the SKOS compliant Semantic Web RESTful services. Note that in the present study, we did not implement the CTS2 model as illustrated in the system architecture, instead we evaluated the potential of the SKOS-based approach in support of building the CTS2 compliant services for the ICD-11 linked data in the next section.

4.1 Normalize the ICD-11 Data Using a SKOS-based RDF Model

The first part of the normalization is to generate the unique identifier for each ICD-11 foundation entity. In this study, we utilized the uniform entity IDs generated by the WHO. For example, for the entity "Acute myocardial infarction" in the current ICD OWL-based rendering, the original ICD-10 code I21 is used as the entity ID, and the URI of the entity is <u>http://who.int/icd#I21</u>. The uniform entity IDs 1634725920 is generated by the WHO for this entity.

The second part of the normalization is to implement the ICD URI scheme proposed by the WHO. For example, the normalized URI for the entity "Acute myocardial infarction" becomes <u>http://id.who.int/icd/entity/1634725920</u>.

The third part of the normalization is to utilize the Simple Knowledge Organization System (SKOS) RDF data model. In the current prototype implementation, we defined the mappings between the ICD-11 model signatures and the SKOS model signatures (Table 1). Figure 3 shows a normalized ICD-11 entity "Acute myocardial infarction" using the SKOS data model.

Table 1. The mappings between the ICD-11 model signatures and the SKOS model signatures.

ICD-11 Predicate	SKOS Predicate
	skos:ConceptScheme,
ICD Classification	skos:inScheme
ICD Entity	skos:Concept
icd:entityID	skos:notation
icd:icdTitle	skos:prefLabel
icd:synonym, icd:fullySpecifiedName	skos:altLabel
icd:definition	skos:definition
icd:parent, owl:subClassOf	skos:broaderTransitive
icd:child, owl:subClassOf(inverse)	skos:narrowerTransitive

icd - http://id.who.int/icd/schema/

skos - http://www.w3.org/2004/02/skos/core#/



Figure 3. A RDF Turtle rendering of a normalized ICD-11 entity "Acute myocardial informatics" using the SKOS data model. Note that the text is wrapped for the display purpose. Note that a SNOMED CT concept 74281007 (Myocardium structure) is used as an ontological element for describing the "bodyPart" of the ICD-11 entity.

4.2 Build Semantic Web RESTful Services for the ICD-11 Linked Data

We used the Elda implementation of Linked Data API. We created an API specification file as required by the Elda. Basically we specified the following modules for generating a RESTful endpoint as a series of steps. First, we defined two URI templates that identified the set of URLs to which it should respond; one is a List Template that selects all ICD foundation entities and the other is an Item Template that selects a single ICD entity resource. Second, we specified a number of Selectors that describe how to retrieve a list of resources from the configured SPARQL endpoint.

Here, we use the SPARQL endpoint enabled by a 4store built-in SPARQL server. The 4store is an open source RDF store developed at the Garlik [16]. By defining the Selectors, the RESTful endpoint actually converted them to the corresponding SPARQL queries to retrieve the desired resources. Third, we defined a set of named Viewers that describe specific sets of properties to retrieve about each resource. There is a default viewer that can retrieve all properties available for a specific category. In the prototype implementation, we also defined three customized viewers, the first one called "label viewer" that shows the skos:notation and skos:prefLabel for a specific entity; the second one called "definition viewer" that shows the definition of an entity in addition to those defined in the "label viewer"; the third one called "hierarchy viewer" that shows the children and parent of an entity in addition to those defined in the "label viewer"; the default named Formatters that define the possible representation formats that can be delivered to a client. Table 2 shows the RESTful APIs for an ICD entity "1634725920", its definition view and available formats.

Using the Elda implementation of Linked Data API, we successfully established a RESTful endpoint over the ICD-11 Linked Data. The RESTful endpoint can be accessed as either a List Endpoint or an Item Endpoint. For the List Endpoint, it will provide the RDF graph data for all ICD entities with pagination using the following URL: {baseURL}/icd?_page=0. For the Item Endpoint, it will provide a RDF graph data for a single ICD entity resource. An example URL is. {baseURL}/icd/entity/1634725920. To access the renderings in different formats such as json, xml, rdf, ttl, etc., end users just need to specify the URLs as illustrated in Table 2.

RESTful APIs	Description	
{baseURL}/icd/entity/1634725920?_view=definition	default as html in the view of definition	
{baseURL}/icd/entity/1634725920.json?_view=definition	return json format	
{baseURL}/icd/entity/1634725920.xml?_view=definition	return xml format	
{baseURL}/icd/entity/1634725920.rdf?_view=definition	return rdf format	
{baseURL}/icd/entity/1634725920.ttl?_view=definition	return turtle format	
{baseURL}/icd/entity/1634725920.text?_view=definition	return plain text format	
{baseURL}/icd/entity/1634725920.plainhtml?_view=definition	return plain html format	
{baseURL}/icd/entity/1634725920.html?_view=definition	return html format	

Table 2. The example of RESTful APIs for an ICD entity "1634725920"

5 Preliminary Evaluation

The ultimate goal of building the standardized Semantic Web RESTful services on the ICD-11 linked data is to support the authoring application development of the ICD-11. We believe the usefulness of the system can be evaluated in the following two aspects: 1) to test whether it is feasible to build AJAX (i.e. an acronym for Asynchronous JavaScript and XML) widgets utilizing the RESTful services developed in this study; and 2) to evaluate the potential for extending the SKOSbased approach in support of building CTS2 compliant services on the ICD-11 linked data.

For the first aspect, we invited two experienced web developers at Mayo Clinic to evaluate the feasibility of the services to support the AJAX widget development. A subjective 1-5 scale survey is used for both the feasibility evaluation, with 1 indicating "strongly disagree" and 5 indicating "strongly agree". As the results, both of web developers strongly agreed (i.e. scored as 5) that it is feasible to utilize the RESTful services to build the AJAX widgets in support of collaborative authoring of ICD-11. In addition, we are working on testing the services using a real world AJAX application - an ICD Proposal Authoring Widget based on the SmartGWT (developed in our ongoing study at Mayo Clinic[17]). The widget will interface with the standardized Semantic Web RESTful services developed in the present study for the ICD-11 content access. Figure 4 shows a screenshot of the ICD Proposal Authoring Widget.

http://who.int/icd#l21		
Expand	New Child Concept Multicat	egory Propos
roperties	Details	Action
ARENTS		٢
Parent	Tabulated - Acute myocardial infarction	2
Parent	Ischaemic heart diseases	2
DEFINITION		
O Textual Definitions		٢
Textual Definition	Wyocardial infraction (MI) is defined as of heart muscle cells. Wyocardial infraction occurs due to the lack of oxygen which maintains normal operating function. This generally happens due to interruption of blood supply of the heart when a coronary artery is blockad. Underlying this blockage is generally a collection of lipids and atherosciencic plaques.	S 🕹
ICD Title	Acute myocardial infarction	N
Fully Specified Name		N
🛯 🎱 TERMS		
CLINICAL DESCRIPTION		
I 🌍 DIAGNOSTIC CRITERIA		
🛯 🌍 CASUAL MECHANISMS AND RISK FACTORS		
I 🌍 FUNCTIONAL IMPACT		
	I	

Figure 4. A screenshot of an ICD Proposal Authoring Widget based on the SmartGWT.

For the second aspect, we tried to define the mappings between the SKOS model predicates and the CTS2 model predicates. Since the SKOS data model is a proper subset of the CTS2 model, we were able to define the mappings between the SKOS

model predicates and the CTS2 model predicates. Table 3 shows the mappings between CTS2 predicates and SKOS predicates for this preliminary evaluation. The CTS2 predicate is defined in CTS2 entity description model [13].

CTS2 Predicate	SKOS Predicate
cts2:about	rdf:about
cts2:entityID	icd11:(skos:notation)
cts2:alternateEntityID	-
cts2:describingCodeSystemVersion	skos:inScheme
	skos:prefLabel,
	skos:altLabel,
cts2:designation	skos:hiddenLabel
cts2:definition	skos:definition
cts2:note	skos:note
cts2:parent	skos:broaderTransitive
cts2:children	skos:narrowerTransitive

Table 3. The mappings between CTS2 predicates and SKOS predicates

cts2 - <u>http://id.omg.org/spec/cts2#</u>

icd11 - http://id.who.int/icd/entity/

skos - http://www.w3.org/2004/02/skos/core#/

rdfs - http://www.w3.org/2000/01/rdf-schema#

rdf -http://www.w3.org/1999/02/22-rdf-syntax-ns#

6 Discussion

In this study, we developed a framework of building standardized Semantic RESTful services on ICD-11 linked data that are transformed from an ICD-11 relational database. In the prototype implementation of the framework, we have utilized a number of open source Semantic Web applications: Protégé OWL/RDF exporter, 4store RDF store and Elda implementation of Linked Data API. As shown in the semantic normalization layer in Figure 3, we used the Protégé OWL/RDF exporter to generate an OWL/RDF rendering of ICD-11 contents for subsequent normalization. The 4store RDF store enabled the storage of ICD-11 linked data and a built-in SPARQL endpoint server. The Elda Linked Data API provides a highly configurable mechanism that allowed us to build the simple RESTful services on top of a SPARQL endpoint.

The original rationale of the Linked Data API is that simple RESTful APIs are well supported and understood by a large community of web developers and SPARQL is a power tool whose sophistication is unnecessary for many users [14]. The goal of the API is to provide some basic functionality "out-of-the-box", but provide a range of configuration options to support additional, domain-specific customization and functionality. Basically, the Linked Data API supports encoding the requests from the end users in a API specification file and the underlying engine will convert the specification into the corresponding SPARQL queries, retrieve back the requested data in RDF graph and expose the data as the RESTful services in a number of easyto-process representation of resources such as simple JSON, or XML, etc. In the biomedical domains, for instance, the similar approach was used to build a lightweight web service interface for Semantic Web contents integrating multiple life science databases in a Semantic-JSON project [18].

Combining the strengths of the Semantic Web application components above, we actually developed an open-data solution that would be useful to maximize the data accessibility to support the authoring application development of the ICD-11 for the beta phase. This solution has made the infrastructure of ICD-11 data available as the RESTful web services such that the ICD-11 linked data are pretty friendly to be consumed by many existing AJAX or AJAX-like frameworks and toolkits. The Exhibit, for example, is a lightweight framework for publishing structured data on standard Web servers that requires no installation, database administration, or programming [19]. Exhibit has been used widely in many web applications and projects such as a patient record data collection project [20], a collaborative platform for representation and harmonization of clinical data elements [21], etc. The Google Web Toolkit (GWT) is another example of such an AJAX-based framework that has been used to develop the WHO ICD-11 alpha authoring tool [3-4]. In addition, WHO is currently using a .NET-based browser to expose ICD-11 alpha content [22]. The browser provides simple commenting functionality to allow the domain professionals to make comments on existing contents, and it intends to introduce more social computing capabilities. As the browser is actually using some AJAX widgets, we consider that the browser should be able to consume the ICD-11 contents through the standardized services developed in the present study. As part of our preliminary evaluation, we evaluated the feasibility of the RESTful services developed in this study in support of the AJAX widgets. Two experienced web developers all strongly agreed that it is highly feasible to build the AJAX widgets by utilizing the RESTful services.

We have made efforts on normalizing the ICD-11 linked data using the SKOS RDF model along with the WHO ICD URI scheme. The SKOS model provides a set of Semantic Web friendly signatures with well-defined semantics. For instance, the SKOS not only provides a rich set of labeling and document properties (e.g., skos:prefLabel, skos:altLabel, skos: note, skos:definition, etc), but also provides welldifferentiated hierarchical semantic relationships, e.g., differentiating between skos:broader and skos:broaderTransitive and between skos:narrower and skos: narrowerTransitive. We consider that the use of the SKOS model would facilitate the access and use of ICD-11 classification by many existing SKOS-aware Semantic Web tools. In addition, the SKOS RDF data model is defined as an OWL full ontology, so it is highly expressive and extensible. We consider that this data model is particularly suitable for representing the foundation component in the ICD-11, which is based on a highly complex model involving many complicated use cases. Along with the SKOS RDF model, we have implemented the ICD URI scheme proposed by WHO. Note that the URI scheme is under active discussion in the WHO ICD revision community. We consider that this implementation played a role in examining the proposed URI scheme. For example, in the SKOS model, an ICD entity is represented as an instance of skos:concept, in which a property skos:inScheme is used to specify the context information about concept scheme. We found that the URI for naming such concept scheme is required though original proposal did not cover this requirement.

In addition, we have made efforts in evaluating the potential of the ICD-11 SKOS model in support of building the CTS2 compliant Semantic Web RESTful services. We envision that the CTS2 standard will provide additional layer of standardization that would potentially optimize the interoperability between downstream authoring applications. The goal of the CTS2 standard is to provide a common model and semantics for the representation, interchange and federation of terminological resources. For the ICD-11 revision use case, part of the input process involves a way for users to download, test and evaluate ICD-11 in a variety of situations, including its alignment with SNOMED CT and other terminologies. Another component of that process involves the ability to propose additions and changes – something that, if successful, will need to be managed in a semi-automated fashion. The key is, in part, a public, open-data model and associated contents that cover not just ICD-11 but ICD-10, SNOMED CT, ICD and others, and allow proposed changes to be accepted, reviewed, revised and adopted or rejected in a semi-automated fashion. The emerging CTS2 standard embodies many of these characteristics already, which is why our study evaluates it as a possible solution. One of CTS2 efforts is on defining a canonical RDF specification that establishes the RDF tags and structure to represent the various CTS2 resources in RDF format. The SKOS rendering in this paper is considerably close to this target. In the preliminary evaluation, we have identified an initial set of mappings between the SKOS predicates and the CTS2 signatures. We consider the mappings would serve as a starting point in support of building the CTS2 compliant services in the future.

There are a number of limitations in this study. Firstly, the mappings among ICD-11 content model, SKOS data model, CTS2 standard and have not yet been extensively discussed in the ICD-11 community, although we consider the draft mappings presented in this study are accurate in certain degree as the authors have been involved in the specifications of these models. Secondly, although the preliminary evaluation results are promising, it seems clear that we could obtain more reliable results by inviting more web developers and CTS2 experts for the evaluation. Thirdly, the RESTful services developed in this study are limited to a number of common use cases. A systematic review and collection of the use cases and requirements from the ICD-11 revision community would be helpful to make the services more useful to the broader parties in the community.

In summary, we described our experiences and challenges on building the SKOS compliant Semantic Web RESTful services that aim to support the collaborative authoring application development of ICD-11. We have demonstrated the usefulness of the proposed framework in terms of its feasibility and its potential in support of building CTS2 compliant services. In the future studies, we plan to integrate CTS2 standard to consolidate and standardize the RESTful services based on the evaluation from the ICD-11 community and to build reusable AJAX widgets library utilizing the ICD-11 RESTful services. We expect our efforts in the present study would facilitate data accessibility and system interoperability that have been challenging in the ICD-11 community.

Acknowledgement

The authors would like to thank two experienced web developers: Kevin Peterson and Cory Endle and two CTS2 experts: Craig Stancl, Harold Solbrig, who participated in the evaluation of the study.

References

- 1. WHO. Revision of the International Classification of Diseases (ICD). Available from:
- http://www.who.int/classifications/icd/ICDRevision/en/index.html.
 Chute CG. Distributed biomedical terminology development: from experiments to open process. Yearb Med Inform. 2010:58-63.
- T. Tudorache, C. I. Nyulas, N. F. Noy, T. Redmond, M. A. Musen. iCAT: A Collaborative Authoring Tool for ICD-11. Workshop on Ontologies come of Age in the Semantic Web at ISCW 2011, Bonn, Germany. Published in 2011.
- Tudorache T, Falconer S, Nyulas C, Storey MA, Ustün TB, Musen MA. Supporting the Collaborative Authoring of ICD-11 with WebProtégé. AMIA Annu Symp Proc. 2010 Nov 13;2010:802-6.
- WHO Software Development Meeting 2011: http://sites.google.com/site/icd11revision/home/face-to-facemeetings/software-development-meeting-11-april-2011; last visited at September 28, 2011.
- Samwald M, Jentzsch A, Bouton C, Kallesøe CS, Willighagen E, Hajagos J, Marshall MS, Prud'hommeaux E, Hassanzadeh O, Pichler E and Stephens S. Linked open drug data for pharmaceutical research and development. Journal of Cheminformatics 2011, 3:19.
- 7. BioPortal: http://bioportal.bioontology.org/; last visited September 28, 2011.
- Jiang G, Solbrig H, Chute CG. ADEpedia: A scalable and standardized knowledge base of adverse drug events using Semantic Web technology. AMIA Annu Symp Proc. 2011;2011:607-16. Epub 2011 Oct 22.
- 9. SKOS Reference: http://www.w3.org/TR/2009/REC-skos-reference-20090818/. Last visited at October 1, 2012.
- 10. ICD-11 Content Models in UML: http://informatics.mayo.edu/icdmodel. Last visited at September 27, 2012.
- 11. Agreement between IHTSDO and WHO URL: http://www.who.int/classifications/AnnouncementLetter.pdf. Last visited at August 8, 2012.
- 12. Linked Data Design issues: http://www.w3.org/DesignIssues/LinkedData.html; last visited at September 27, 2012.
- 13. CTS2 wiki: http://informatics.mayo.edu/cts2. Last visited at September 28, 2012.
- 14. Linked Data API: http://code.google.com/p/linked-data-api/; last visited at September 28, 2012.

- 15. Elda implementation of Linkded Data API: http://code.google.com/p/elda/; last visited at September 28, 2012.
- 16. 4store RDF store: http://4store.org. last visited at September 27, 2012.
- 17. Jiang G, Solbrig HR, Chute CG. Using semantic web technology to support ICD-11 textual definitions authoring. ACM International Conference Proceeding Series. 2012; 38-44.
- Kobayashi N, Ishii M, Takahashi S, Mochizuki Y, Matsushima A, Toyoda T. Semantic-JSON: a lightweight web service interface for Semantic Web contents integrating multiple life science databases. Nucleic Acids Res. 2011 Jul;39(Web Server issue):W533-40. Epub 2011 Jun 1.
- 19. Huynh D, Karger D and Miller R. Exhibit: Lightweight structured data publishing. In Proceedings of the 16th international conference on World Web Web; page 746. ACM, 2007.
- Ogbuji C. A role for Semantic Web technologies in patient record data collection. In the book – Linking Enterprise Data (Wood D, Ed.), 2010 – ISBN 978-1-4419-7664-2.
- Jiang G, Solbrig HR, Iberson-Hurst D, Kush RD, Chute CG. A Collaborative Framework for Representation and Harmonization of Clinical Study Data Elements Using Semantic MediaWiki. AMIA Summits Transl Sci Proc. 2010 Mar 1;2010:11-5.
- 22. WHO ICD-11 Beta Browser: http://apps.who.int/classifications/icd11/browse/f/en. Last visited at October 1, 2012.