e-Health – For Continuity of Care
C. Lovis et al. (Eds.)

doi:10.3233/978-1-61499-432-9-1043

ICD-11 and SNOMED CT Common Ontology: Circulatory System

Jean-Marie RODRIGUES[a,b], Stefan SCHULZ[c], Alan RECTOR[D], Kent SPACKMAN[e], Jane MILLAR[e], James CAMPBELL[f], Bedirhan ÜSTÜN[g], Christopher G. CHUTE [h], Harold SOLBRIG[h], Vincenzo Della MEA[i], Kristina Brand PERSSON[j]

[a] INSERM U1142, LIMICS, Paris, France
[b] *Department of Public Health and Medical Informatics, Univ. of Saint Etienne, France*
[c] *IMI, Medical University of Graz, Austria*
[d] *University of Manchester, UK*
[e] *Intl Health Terminology Standards Development Organisation, Copenhagen*
[f] University of Nebraska Medical Center Omaha USA
[g] *World Health Organization, Geneva, Switzerland*
[h] *Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, USA*
[i] *Department of Mathematics and Computer Science, University of Udine, Italy*
[j] *National Board of Health and Welfare, Stockholm, Sweden*

Abstract. The improvement of semantic interoperability between data in electronic health records and aggregated data for health statistics requires efforts to carefully align the two domain terminologies ICD and SNOMED CT. Both represent a new generation of ontology-based terminologies and classifications. The proposed alignment of these two systems and, in consequence, the validity of their cross-utilisation, requires a specific resource, named Common Ontology. We present the ICD-11 SNOMED CT Common Ontology building process including: a) the principles proposed for aligning the two systems with the help of a common model of meaning, b) the design of this common ontology, and c) preliminary results of the application to the diseases of the circulatory system.

Keywords. ICD, SNOMED CT, Ontology, Terminology, Classification

Introduction

The increasing availability of machine-readable data in medicine, health care and public health has put the desideratum of **Semantic Interoperability** on the top of the Health Informatics agenda. It targets the preservation of meaning between heterogeneous patient-related and aggregated population data across different vocabularies and coding systems. This coincides with the current revision process of the disease classification system ICD [1], which continues being the most important semantic standard for health statistics, but which is also used – in several national modifications and extensions – for morbidity documentation and billing. This occurs in a time in which the international clinical terminology standard SNOMED CT (SCT) [2]

[3] has gradually been gaining ground. SCT, maintained by the IHTSDO as an international medical terminology standard, promises to provide codes and formalisms to represent the whole health care process.

Clinicians, documentation specialists, epidemiologists and health care administrators may rightly ask why they should use SCT in parallel to ICD and local procedure classification systems. This concern was one of the drivers for a collaboration agreement between the maintainers of ICD and SCT, *viz.* WHO and IHTSDO. This agreement has fuelled the identification of a common basis between both terminological systems and culminated in the decision to create a **Common Ontology (CO)** [4]. We here report on the current state of this common ontology.

1. Materials and Methods

1.1. ICD-11 – SNOMED CT Harmonization

In 2007, the WHO launched the revision of the International Classification of Diseases – ICD-11 [5]. This revision process is different from previous ones in that the authoring is computerized and supported by ontology-driven tools [6] and distinguishes between a multi-hierarchical ICD Foundation Component (FC), from which multiple purpose-specific mono-hierarchical Linearizations can be derived.

After the WHO – IHTSDO agreement and the establishment of a Joint Advisory Group (JAG), it was decided that the *Foundation Component* is intended to have at its core a common ontology which is a subset of SCT, which is an example of bringing a medical ontology into practice. This is in line with the tendency of a new generation of biomedical vocabularies to be built on the principles of applied ontology [7][8][9] [10] [11] and a model of meaning based on descriptive logic [12], using formalisms and languages from the Semantic Web community like OWL [13]. Due to the broad scope and a high degree of granularity of SCT it was decided that for initial testing, a subset would be sufficient to play the role of test set, provided that some gaps (e.g. rare diseases) be filled. We chose the diseases of the circulatory system for evaluation.

As the source from which the common ontology should be drawn the SCT hierarchies *Clinical Findings*, *Disorders* had been identified. The JAG collected evidence that these concepts denote, first of all, clinical situations, i.e. phases of a patient's life in which a given condition of clinical relevance is present [14].

2. Results

The architecture of the coordination between ICD-11 and SNOMED-CT based on the common ontology is illustrated in fig. 1.

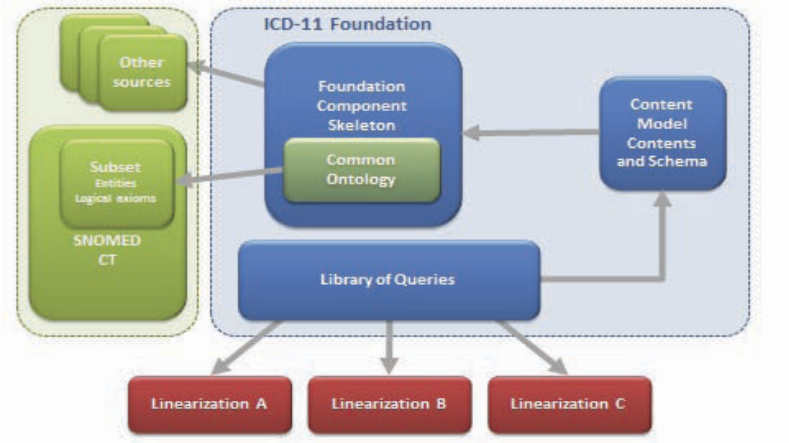

Fig. 1 Architecture of Common Ontology based ICD-11 SNOMED CT coordination

Concisely, it is characterised by the following:

- The linearizations for mortality and morbidity or for primary care are sets of exclusive or disjoint and exhaustive classes meaning that each class is present only once giving a mono-hierarchical look as in ICD-10. It is enforced by classification rules (exclusion statements and residual categories ("not elsewhere classified"; "not otherwise specified")
- The FC is organised around the common ontology coming from a subset of SCT and further on from external ontology sources. This skeleton is populated by the content of all linearizations structured by the common ontology and therefore having multi-hierarchical views. There are two additional elements to this skeleton: the content model and a queries library of SQL style queries needed to relate the linearizations statistical classes with their exclusion and residual rules which cannot be interpreted in strictly logical terms to the CO by preferring "negation as failure" over logical negation.
- The Common Ontology is a subset of SCT classes and logical axioms, from SCT clinical findings and disorder branches. This subset is defined by identifying SCT concept candidates to provide a full definition of the classes of the FC. As in SCT, the Common Ontology is multi-hierarchical.

The workflow is the following, iterating over each ICD-11-FC class:

1. From ICD-11 FC classes as defined on ICD11 browser [15] identify candidate SCT concepts from which the meaning of FC classes can be the same or reconstructed as queries.
   a. Examine the ability of the SCT model of meaning to represent the ICD-11 textual definitions.
   b. Try to fully define all SCT primitive expressions identified in this step and submit for SCT editorial validation.
2. Construct valid SCT formal expressions to reflect the meaning of the ICD-11 textual definitions, to be pre-coordinated in SCT.

3. Specify the queries for ICD-11 FC classes in Boolean constructs of queries against SCT axioms and assess for comprehensiveness, consistency and accuracy.
4. Assure that all expressions in SCT representing subchapters in ICD-11 aggregate to the SCT pre-coordinated expression for the ICD-11 chapter concept. When necessary, submit potential revisions to the SCT concepts model and expansions of SCT content for editorial revision and detailed annotation for future reference.
5. Add all SCT editorial expansions into the common ontology.

2.1. Application to the Circulatory System Disease use case

The application of the method to the circulatory system chapter of ICD-11 has been finalised for steps 1 and 2 for most entities. It is ongoing for ENT, Respiratory, Infectious diseases, Dermatology and Gynaecology and Obstetrics: they concern 22 % of ICD-11 beta draft entities.

The overall results show that among 1582 ICD-11 FC circulatory chapter entities 1412 (89.3%) can be represented by the SCT model of meaning either directly or through a pre-coordinated/post –coordinated alignment between the two systems. 114 (7.2%) are residuals which have to be cleaned from the FC and 56 (3.5%) needs a revision of SCT formal model of meaning. We propose two contradictory examples:

1. The defining relationships of SCT allow the full representation of both the ICD-11 FC class and the SCT concept. An example is FC *Coronary vaso-spastic disease with angina* is necessarily and sufficiently defined as SCT 87343002 *Prinzmetal angina (disorder)*.
2. The SCT defining relationships are not providing a complete ontological representation of both entities, and the concept model or pre-coordinated SCT content must be expanded. An Example is ICD-11 *Acute myocardial infarction, STEMI anterior wall and SCT 401303003 |Acute ST segment elevation myocardial infarction (disorder) and SCT 54329005 | Acute anterior myocardial infarction.*

3. Conclusion

What the ICD-11 SCT common ontology proposes is definitely not the development of yet another ontology. Instead, the goal is to define an applied ontology refurbished from a SCT subset common to on the one hand, ICD-11, the semantic standard for health statistics in mortality, morbidity, documentation and billing, and on the other hand, SCT, the most fine grained medical terminology system.

This will require certain refinement / redesign efforts on both the IHTSDO and WHO side. The test on the ICD-11 beta draft circulatory chapter shows that this could affect around 15 % of the ICD-11. The maintenance will be shared by WHO and IHTSDO. New ontology sources will have to be taken in consideration.

Further on, this FC with the common ontology shall be used for the maintenance of all of the existing WHO ICD as well as the ICD-(10/11) national modifications thereby easing international comparisons and backward compatibility with current systems.

Acknowledgments

This work was supported by the World Health Organization (WHO) and the International Health Terminology Standards Development Organization (IHTSDO) through their Joint Advisory Group (JAG).

References

[1] The World Health Organization, The International Classification of Diseases 11th Revision is due by 2017, http://www.who.int/classifications/icd/revision/en/ (last accessed 29 Apr 2013).
[2] International Health Terminology Standards Development Organisation (IHTSDO): SNOMED CT: http://www.ihtsdo.org/snomed-ct/ (last accessed 20 Jan 2014).
[3] SNOMED CT browser: https:/uts.nlm.nih.gov/ (last accessed 28 Apr 2014)
[4] J.M.Rodrigues, S.Schulz, A.Rector, K.Spackman, T.B.Üstün, C.G.Chute, V.Della Mea, J.Millar, K.Brand Persson, Sharing Ontology between ICD 11 and SNOMED CT will enable seamless re-use and semantic interoperability. Studies in Health Technology and Informatics 2013;192:343-346.
[5] J.M.Rodrigues, A.Kumar, C.Bousquet, B.Trombert, Using the CEN/ISO standard for categorial structure to harmonise the development of WHO international terminologies, *Studies in Health Technology and Informatics* **150** (2009), 255-259.
[6] T.Tudorache, S.Falconer, C.Nyulas, M.A.Storey, T.B.Ustün, M.A.Musen, Supporting the Collaborative Authoring of ICD-11 with Web Protégé, Proceedings of the AMIA Annual Symposium 2010, 802-806.
[7] B.Smith, Applied Ontology: A New Discipline is Born, *Philosophy Today*, **29**(12) (1998), 5-6.
[8] M.Ashburner et al. Gene Ontology: tool for the unification of biology. *Nature Genetics* **25**(2000), 25:29.
[9] OBO foundry principles: http://obofoundry.org/ontologies.shtml (last accessed 5 Feb 2014).
[10] A.Rector, A.Rossi, M.F. Consorti, P.Zanstra, Practical development of re-usable terminologies: GALEN-IN-USE and the GALEN Organisation, *International Journal of Medical Informatics* **48** (1998), 71-84.
[11] S.Schulz, R.Cornet, K.Spackman, Consolidating SNOMED CT's ontological commitment, *Applied Ontology* **6** (2011), 1-11.
[12] F.Baader et al. The Description Logic Handbook: Theory, Implementation, and Applications, 2nd Edition. Cambridge University Press, 2nd ed., 2007.
[13] W3C OWL Working Group: OWL 2 Web Ontology Language Document Overview 2009, http://www.w3.org/TR/owl2-overview/ (last accessed 28 Apr 2014).
[14] S.Schulz, A.Rector, J.M.Rodrigues, K.Spackman, Competing interpretations of disorder codes in SNOMED CT and ICD. AMIA Annu Symp Proc. 2012, 819-827.
[15] The World Health Organization, ICD11 beta draft: http://www.who.int/classifications/icd11 (last accessed 28 Apr 2014)