Medical Informatics in a United and Healthy Europe
K.-P. Adlassnig et al. (Eds.)
IOS Press, 2009

doi:10.3233/978-1-60750-044-5-255

Using the CEN/ISO Standard for Categorial Structure to Harmonise the Development of WHO International Terminologies

Jean M. RODRIGUES [a,b,1], Anand KUMAR [a], Cédric BOUSQUET [a], Béatrice TROMBERT [a]
[a] *Department of Public Health and Medical Informatics, University Jean Monnet, Saint Etienne, France*
[b] *WHO FIC Collaborative Centre for French Language, Paris, France*

Abstract. Semantic interoperability (SIOp) is a major issue for health care systems having to share information across professionals, teams, legacies, countries, languages and citizens. The World Health Organisation (WHO) develops and updates a family of health care terminologies (ICD, ICF, ICHI and ICPS) and has embarked on an open web-based cooperation to revise ICD 11 using ontology driven tools. The International Health Terminology Standard Development Organisation (IHTSDO) updates, translates and maps SNOMED CT to ICD 10. We present the application of the CEN/ISO standard on categorial structure to bind terminologies and ontologies to harmonise and to map between these international terminologies.

Keywords. standard, international terminologies, categorial structure, ontology, interoperability

1. Introduction

Content interoperability has now emerged in health care as a top challenge under the name of semantic interoperability. There is an increasing need to address national and international comparisons, sharing and cooperation across professionals, teams, legacies, languages and citizens for population-based WHO indicators, Electronic Health Record safety, trans-border migration of population, case mix and procedure payment, et al.

Unfortunately clinical terminological systems, classifications and coding systems have been developed by independent, divergent and uncoordinated approaches which have produced non reusable systems on overlapping fields for different needs: WHO-FIC, International Classification of Diseases (ICD), International Classification of Function (ICF), International Classification of Health Intervention (ICHI) and International Classification of Patient Safety (ICPS) [1], UMLS (Unified Medical Language System) [2], LOINC [3] for clinical laboratories, DICOM SDM [4] for

[1] Corresponding Author: Jean M Rodrigues, SSPIM/CIM42, CHU Saint Etienne, Chemin de la marandière, 42270 Saint Priest en Jarez, France; E-mail: rodrigues@univ-st-etienne.fr.

imaging, SNOMED CT [5]. Finally a lot of interface or point of care terminologies are used as the Convergent Medical Terminology (CMT) [6].

Natural language expressions show inconsistencies and ambiguities as it can be assessed by biomedical ontology driven tools [7–9]. The most important achievements are GALEN (Generalised Architecture for Languages, Encyclopaedias and Nomenclatures in Medicine) [10], FMA (Foundational Model of Anatomy) [11–13] and OBO (Open Biomedical Ontology) [14]. WHO has initiated the revision of ICD which will take advantage of such achievements to enhance the interoperability of products of the family of international classifications by the year 2015.

We present in section 2 the new organisation of the revision by the WHO Family of International classifications (WHO FIC) of ICD 11. In section 3 we show the use of the categorial structure standard approach elaborated and developed by the European Standard Body CEN and ISO [15] to bind terminologies to ontologies to harmonise the future developments of the international terminologies. We finally discuss the role of this approach to insure an important shift from divergent systems to a common interoperable and coordinated community.

2. WHO ICD 11 Revision

Historically, ICD evolved to serve international comparisons of mortality. Over the decades, particularly in the latter half of the 20^{th} century, demands for morbidity information drove expansion and application of the ICD. With this increasing need, corresponding uses of ICD for reimbursement and clinical quality were established. WHO has decided [1] that the revision process of ICD 11 will take care of these different use cases: mortality (certifying the death, coding the causes of death, selecting the underlying cause of death)**,** morbidity (hospital statistics, epidemiology and public health including reimbursement based on case mix), primary care (first level of care where diagnostic picture is not fully developed or investigated), quality and safety of care and scientific consensus on clinical phenotypes to distinguish its use for diagnosis and definition of diseases.

A new ICD 11 revision editorial process very different from the previous ones has been put in place. First it will be carried out on a cooperative web-based joint authoring platform instead of annual revision conferences. Second the informatics infrastructure enabling multiple working groups of clinicians to make their contributions to the revision process in a standard fashion will be based on a content model and on templates that will be used by contributors from the different clinical colleges through well-defined value sets using terms from predefined terminologies. Finally ICD 11 will have explicit linkages to underpinning ontologies and will result in both human-readable definitions for end users and machine-readable definitions for automatic retrieval, translating, mapping and aggregation.

3. CEN/ISO Categorial Structure Standard Approach for ICD 11 Revision

3.1. CEN/ISO Categorial Structures

We have presented elsewhere the history, rationale and contents of these standards [15, 16]. The CEN/ISO Categorial structure was defined within some linguistic variations [15] as a minimal set of health care domain constraints to represent a biomedical terminology (controlled vocabularies, nomenclatures, coding systems and classifications) in a precise health care domain with a precise goal to communicate safely. The categorial structure proposes a frame for a lite ontological organisation to ensure standardisation of the knowledge representation of terminologies a way to bind terminology with ontology without description logic.

3.2. ICD 11 Revision Categorial Structure Role

The ICD 11 Revision process will use a cooperative web-based joint authoring platform based on a content model and on templates for clinical contributors. The contributors will be medical domains experts from the different clinical colleges related to a specific domain Topic Advisory Group (topic TAG) as internal medicine, mental health, rare diseases, etc. The finalized content model shall be aligned with the researches and development in ontology and description logic using the Protégé editor [17, 18].

3.3. ICD 11 Revision Categorial Structure Content

3.3.1. First Distinction

It is necessary to do a distinction between the pre-coordinated concepts categories or the axis of the terminology as Human Anatomy, Body Function, Morphology, Cause, Severity, Occurrence, Stage etc. and the post-coordinated categories which is the association of the previous categories to represent the knowledge (Figure 1).

3.3.2. Concerning the Post Coordinated Categories

There is enough evidence and namely with the different uses cases identified that at least three approaches shall be considered related between them in an architecture of Russian dolls and can be called provisionally: Disease, Diagnosis, Patient Findings and Problems.

Disease is the most complete view as in the mortality use case and in the clinical phenotypes use case. It is based on an abnormality in the body structure (morphology) or in the body function (patho-physiology), a Cause which can be deterministic as environmental or probabilistic as genetic plus the characteristics of the 2 other views included in it.

Diagnosis is the view of a clinical decision maker who shall take a decision in an uncertain situation as referred by the morbidity and quality and safety use cases. It is based on a set of patient findings and problems to be defined by the domain-specific TAGs but is an assumption and not as evidence-based as the previous one. On the other hand it is a frequent situation where the ICD has to be used as pertinently addressed by the representative of the mental disease TAG.

Finally Patient Findings and Problems (signs, symptoms, syndromes, test results, situations, etc.) are very often mentioned in health record for surveillance or other without reaching the level of a diagnosis assumption. This is well addressed in the ICD revision primary care use case and in the ISO reference terminology standard for nursing.

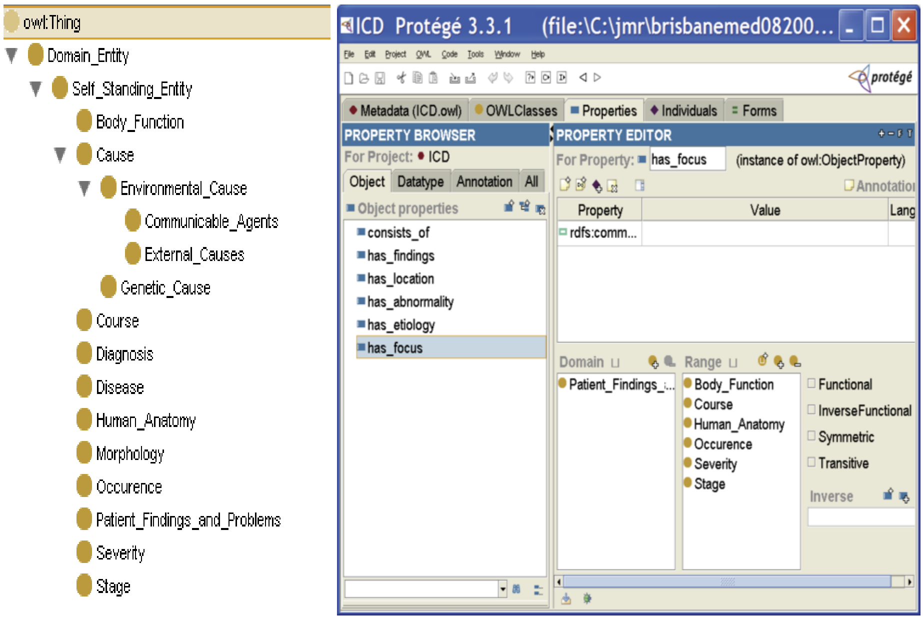

Figure 1. Semantic categories (left) and semantic links (right)

3.3.3. The Semantic Links

They are has_findings, has_location, has_abnormality, has_etiology and has_focus.

- has_findings is the link authorised between Disease or Diagnosis and Patient Findings and Problems
- has_location is the link authorised between Disease or Diagnosis and Human Anatomy
- has_abnormality is the link authorised between Disease and Body Function or Morphology
- has_etiology is the link authorised between Disease and Cause
- has_focus is the link authorised between Patient Findings and Problems and Body Function, Course, Human Anatomy, Occurrence, Severity and Stage.

3.3.4. The Minimal Domain Constraints

Patient Findings and Problems: at least one has_focus
Diagnosis: at least one has_findings and at least one has_location
Disease: at least one has_findings, at least one has_location, at least one has_abnormality and at least one has_etiology

4. Conclusion

We have presented the initial architecture of ICD 11 revision process based on a new organisation and on new tools available in biomedical informatics and in the web open source community. It is an important shift from traditional paper-based revisions restricted to classifications and coding references centres to an unlimited community of contributors allowing to insure the multipurpose uses of ICD 11.

The utilisation of web open source and ontology tools is the guarantee to increase semantic interoperability with other international terminologies within the WHO FIC network (ICF, ICHI and ICPS) and outside (IHTSDO and SNOMED CT). The oral presentation will show practical examples of the uses of the content model categorial structure and of the templates produced by domain-specific revision TAGS.

Acknowledgements. We wish to thank the members of department of Classification, Terminology and Standards from the Health Statistics and Informatics division in the WHO headquarters in Geneva and the coordinator B Ustun, the WHO FIC collaborating centres, our partners in the CEN TC251 WG2 and ISO TC 215 WG 3 and the convenors, terminology and biomedical ontology experts namely Alan Rector from University of Manchester and Barry Smith from the University at Buffalo and our partners in the GALEN program.

References

[1] http://www.who.int/classifications/en/.
[2] McCray, A.T., Nelson, S.J. (1995) The representation of meaning in the UMLS. *Methods of Information in Medicine* 34(1-2):193–201.
[3] http://www.ihtsdo.org/snomed-ct/history0.
[4] Logical Observation Identifiers Names and Codes (LOINC), http://www.loinc.org/.
[5] DICOM. http://www.xray.hmc.psu.edu/dicom/dicom_home.html.
[6] Dolin, R.H. (2003) Kaiser Permanente's Convergent Medical Terminology. Testimony to the National Committee on Vital and Health Statistics, Subcommittee on Standards and Security. May 22, 2003. http://ncvhs.hhs.gov/030522sstr.htm.
[7] Rodrigues, J.-M., Rector, A., Zanstra, P. et al. (2006) An ontology driven collaborative development for biomedical terminologies: From the French CCAM to the Australian ICHI coding system. In Hasman, A. et al. (Eds.) *Proceedings of MIE 2006*, IOS Press, Amsterdam, 863–868.
[8] Gruber, T. (1993) A translation approach to portable ontology specifications. *Knowledge Acquisition* 5(2):199–220.
[9] Smith, B., Ceusters, W. (2003) Towards industrial strength philosophy: How analytical ontology can help medical informatics. *Interdisciplinary Science Reviews* 28:106–111.
[10] Baud, R.H., Rodrigues, J.-M., Wagner, J.C. et al. (1997) Validation of concept representation using natural language generation. In Masys, D.R. (Ed.) *AMIA Annual Symposium Proceedings 1997*, 841.
[11] Rector, A. (1994) Compositional Models of Medical Concepts: Towards Re-usable Application-Independent Medical Terminologies. In Barahona, P., Christensen, J.P. (Eds.) *Knowledge and Decisions in Health Telematics*, IOS Press, Amsterdam, 109–114.
[12] University of Washington Digital Anatomy (UWDA). A reference ontology for biomedical informatics: The Foundational Model of Anatomy (FMA), http://fma.biostr.Washington.edu.
[13] Rosse, C., Mejino, J.V.L. (2003) A reference ontology for biomedical informatics: The Foundational Model of Anatomy. *Journal of Biomedical Informatics* 36:478–500.
[14] Smith, B., Ceusters, W., Klugges, B., Kohler, J., Kumar, A., Limax, J., Mungall, C.J., Neuhaus, F., Rector, A., Rosse, C. (2005) Relations in biomedical ontologies. *Genome Biology* 6: R46.
[15] Rodrigues, J.-M. et al. (2008) Standards and Biomedical Terminologies: The CEN TC 251 and ISO TC 215 Categorial Structures. A Step towards increased interoperability. In Andersen, S.K. et al. (Eds.) *Proceedings of MIE 2008*, IOS Press, Amsterdam, 857–862.
[16] Rodrigues, J.-M., Rosse, C., Fogelberg, M. et al. (2007) A road from health care classifications and coding systems to ontology: The CEN categorial structure for human anatomy: Catanat. In Kuhn, K. et al. (Eds.) *Proceedings of MEDINFO 2007*, IOS Press, Amsterdam, 735–740.
[17] Noy, N.F., Fergerson, R.W., Musen, M.A. (2000) The knowledge model of Protege-2000: Combining interoperability and flexibility. *12th International Conference on Knowledge Engineering and Knowledge Management Methods, Models and Tools (EKAW'2000)*, Springer, Berlin, 69–82.
[18] Haarslev, V., Möller, R. (2001) Description of the RACER System and its Applications. In *Proceedings of the International Workshop on Description Logics (DL-2001)*, Stanford, USA, 131–141.