Using ISO and Semantic Web standards for creating a Multilingual Medical Interface Terminology: A use case for Heart Failure

Elena Cardillo
Fondazione Bruno Kessler, Trento, Italy
cardillo@fbk.eu

Maxime Warnier
Université Catholique de Louvain
Louvain, Belgium
maxime.warnier@student.uclovain.be

Joseph Roumier
CETIC Charleroi, Belgium
joseph.roumier@cetic.be

Marc Jamoulle
IRSS-Université Catholique de Louvain
Louvain, Belgium
marc.jamoulle@uclouvain.be

Robert Vander Stichele
Heymans Institute of Pharmacology
Ghent, Belgium
robert.vanderstichele@ugent.be

To be presented at and published in the Proceedings of the 10th International Conference on Terminology and Artificial Intelligence, Paris, November 28-30, 2013.

Abstract

The correct registration and encoding of medical data in Electronic Health Records is still a major challenge for health care professionals. Efficient terminological systems are lacking to enable multilingual semantic interoperability between general practitioners, patients, medical specialists, and allied health personnel. The aim of this paper is to propose an architectural structure for a Multilingual Medical Interface Terminology. We propose a dual structure with a multilingual reference terminology and a collection of unilingual end-user lexicons. Our methods rely on terminological standards, such as Terminology Markup Framework (ISO 16642) and Lexical Markup Framework (ISO 24613), and on Semantic Web technologies. We present procedures to select words, phrases, and concepts to populate these resources (manual concept extraction, automated term extraction), to link them to NLP applications and international classifications. We present the publication of these resources in Linked Open Data and show the feasibility in a use related to heart failure. We illustrate in particular the difficulties in linking real life concepts (N=168) to multiple international classifications with different functionalities, level of granularity, and scopes. The expansion of entries (from 77 to 298) in the lexicon is shown, when lay term synonyms are considered and decomposing of phrases is performed.

1 Introduction

In spite of the progress reached in the field of Medical Informatics, it is still difficult for Health care professionals and other health actors to register and codify clinical data in daily practice. The link between Electronic Health Records (EHRs) and international English classification systems used to codify the clinical data is often complex, not well integrated and hampered by a translation gap, both in terms of language and of world of reference (patient, general practitioner, medical specialist, etc.).

There is an increased awareness that attaining the goal of semantic interoperability entails construction of interface terminologies, defined by (Rosenbloom et al., 2006) as a ‘‘systematic collection of healthcare-related phrases (terms) to support clinicians’ entries of patient-related information into computer programs such as clinical “note capture” and decision support tools, facilitating display of computer-stored patient information to clinician users as simple human-readable texts’’. These kinds of interface terminologies can be used for problem list entry, clinical documentation in EHRs, text generation and care provider order entry with decision support.
The question is whether existing interface terminologies are sophisticated enough to support semantic interoperable communication of the clinical data between partners in a multilingual health care system (Romary et al., 2006). Until now, interface terminologies have been either limited to one language (often English), or either

providing an interface to only one nomenclature (e.g. SNOMED) or classification (e.g. International Classification of Diseases).

The aim of this paper is to present the architectural structure for a multilingual medical interface terminology, following both lexical and terminological ISO standards on multilingual terminologies and using Semantic Web technologies and languages (RDF/OWL, SPARQL, Linked Data, etc.). The ambition of this multilingual resource for general practitioners, patients, medical professionals, and allied health personnel is to span the gap between human language (as addressed in Natural Language Processing systems) and machine language (used to manage concepts and their lexical representations) and to map to a variety of well-respected international medical nomenclatures, thesauri and classification systems).

The paper is structured as follows: Section 2 gives an overview of the state of the art in the field of Medical Terminology. Section 3 describes the approach to building the structure for a hybrid interface terminology. Section 4 is devoted to present a use case on Heart Failure and preliminary results. Finally, Section 5 provides some discussion and conclusions.

2 Background

Over the last two decades, research on medical terminologies and classification systems has become a popular topic and much work has been done to map between several nomenclatures (UMLS[1], SNOMED-CT[2]), thesauri (MeSH[3]), and classification systems (ICPC[4], ICD[5]), different in structure and purposes and used by physicians during their patients' health care visits.

A number of studies are based, for example, on the extensive use of UMLS as a knowledge resource for extracting semantic mappings (see Fung and Bodenreider, 2005).

Recently, the use of Semantic Web[6] technologies in the biomedical domain has leaded to promising results in terms of information integration across heterogeneous resources. Examples are the use of Resource Development Framework (RDF), triple stores and SPARQL queries in integrating consumer-oriented terminologies with standard classification systems in UMLS (Cardillo et al., 2012), or for aligning standardized nomenclatures with thesauri (Bodenreider, 2008). Many tools have been created for automatic alignment between resources. Two examples are TAXOMAP[7] and ONAGUI[8], more useful to ontology alignment. Results of these kind of alignment algorithms are recorded in file in RDF or OWL[9] language.

The major use of Semantic Web technologies in healthcare is the formalization of existing medical terminologies or classification systems in ontologies. SNOMED-CT (Rector and Brandt, 2008) and the upcoming ICD-11 (Tudorache et al., 2010) represent a good example of this practice. Important is also the creation of medical ontology repositories, such as Bioportal[10], a Web-based open repository where users can search and browse biomedical ontologies as well as create mappings for them.

The use of ontologies in healthcare is very useful, above all if integrated in EHRs or in Personal Health Records (PHRs), since they give structuring and semantics to the recorded information. With predefined mappings of vocabularies used in the original data, they also allow for aggregation, reuse of knowledge, automated reasoning on data, search and retrieval of data from diverse original source systems.

Finally, a number of initiatives have been launched for the creation of consumer-oriented medical vocabularies[11] that map to standardized terminologies or classification systems. Examples of this challenge are: the Open Access Collaborative Consumer Health Vocabulary for English, developed by (Zeng et al., 2006) and available in the UMLS Metathesaurus; the Italian Consumer Medical Vocabulary (ICMV), developed by (Cardillo, 2011) using a lexi-ontological approach, and the Multilingual Glossary of Popular and Technical Medical Terms, in nine European languages[12], developed under initiative of the European Commission.

Crucial missing links in this field are the lack of application in the domain of medical termi-

[1] http://www.nlm.nih.gov/research/umls/
[2] http://www.ihtsdo.org/
[3] http://www.ncbi.nlm.nih.gov/mesh/
[4] http://www.globalfamilydoctor.com/wicc/sensi.html
[5] http://www.who.int/classifications/icd/en/
[6] http://www.w3.org/standards/semanticweb/
[7] https://www.lri.fr/~hamdi/TaxoMap/TaxoMap.html
[8] http://onagui.sourceforge.net/
[9] http://www.w3.org/TR/owl-features/
[10] http://bioportal.bioontology.org/
[11] Lexical resources that reflect the way consumers/patients express and think about health topics.
[12] http://users.ugent.be/~rvdstich/eugloss/information.html

nology of two ISO terminological standards, namely Lexical Markup Framework (LMF, ISO 24613, 2008) and Terminological Markup Framework (TMF, ISO 16642, 2003), the absence of a multilingual approach.

3 Building the architectural structure for a hybrid Interface Terminology

We propose a Medical Interface Terminology, that is composed of two types of domain-specific resources,: a multilingual reference terminology (linked to international classifications, thesauri and nomenclatures), on the one hand, and unilingual end-user lexicons the other hand.

To create a comprehensive interface terminology to be used in the primary domain for information storage, encoding, translation, and retrieval, we used a hybrid approach that consists in the combination between onomasiological and semasiological approaches. On one hand we build the structure of the reference terminology starting from essential concepts of the domain and then look for their lexical representations in different international classifications. On the other hand, we build also the structure of an end-user lexicon following the opposite approach, so starting with a word (or phrase) and looking for its different meanings, and in particular to which concept in the onomasiological resource the word refers to. As mentioned in Section 1., to build this interface terminology we advocate our choice for standard frameworks, using the ISO norms on Terminological Markup Framework (TMF) and on Lexical Markup Framework (LMF), whose meta-models perfectly fit with our requirements. For each type of resource of the interface terminology, we explain our approach to apply the mentioned standards, and to build the two recourses with data categories and domain values from the ISOcat.org platform[13], a data category registry (ISO 12620, 1999). Finally, we describe the approach for publishing these resources as Linked Open Data (LOD).

3.1. Creating a TMF based Reference Terminology

A reference terminology comprehensively and rigorously defines reference concepts and expressions within the biomedical domain, including interrelations between concepts (Rosenbloom et al., 2006), and provides a common reference point for comparisons and aggregation of data about the entire health care process, recorded by multiple different individuals, systems or institutions. It allows the concepts to be defined in a formal and computer-processable way and to be mapped to existing standard nomenclatures and classification systems. This allows to support consistent and understandable coding of clinical concepts and so is a central feature for use in EHRs.

The first resource we created was the multilingual reference terminology, following the standard TMF (Romary et al., 2006), conceived to structurally address multilingualism. The TMF meta-model, which keeps the traditional onomasiological view of a terminological entry, decomposes the organization of a terminological database into five basic components: i) the terminological resource (Terminological Data Collection); ii) the concept (Terminological Entry); iii) the language chosen (Language Selection); iv) the term(s) in the language chosen (Term Entry); v) components of the term (Term Component Section).

The meta-model is ornated with a subset of internationally accepted data categories, relevant to the functionalities of the reference terminology (e.g., *entrySource*, *languageIdentifier*, *preferredTerm*) that have been selected from the already mentioned ISOcat.org platform. Detailed explanations on the TMF meta-model and on the selected data categories can be found in (Roumier et al., 2011).

In a further step, we created a Terminology Markup Language (TML) that comprises the mappings between the Meta-model, the data categories and their domain values. The resulting TML has been serialized using the XML schema language RelaxNG[14]. For this part of the work, we were inspired by the TML created in the TermSciences[15] project (Khayari et al., 2006).

Our aim is to keep the number of core concepts to a respectable minimum (between 7.000 and 15.000 concepts). For each concept, the level of granularity is assessed. Pre-coordinated concepts are kept to a minimum (based on frequency-of-use criteria).

In a first step at the Terminological Entry level (language independent), selected concepts are incorporated in the TMF resource. To each concept a definition in English is assigned and if

[13] http://www.isocat.org/

[14] http://w ww.relaxng.org/

[15] http://www.termsciences.fr/

possible, a perfect match for that concept is looked for in the SNOMED-CT nomenclature, or in the UMLS Metathesaurus. Otherwise, the concept is genuinely defined within the system. Each concept in the reference terminology is categorized according to medical categories (symptom, disease, medical procedure, body part, etc.). In addition, to ensure semantic interoperability, links to international terminologies (SNOMED-CT and UMLS) and classification systems (ICD 10th revision, ICD-10, and ICPC 2nd edition, ICPC-2) are provided, with the corresponding qualitative nature of the mapping (exact match, nearly exact, match to higher or lower level of granularity, match not possible).

In a second step at the Term Entry level (language dependent) the concept is labeled with one preferred term (the most suitable clinical term used by physicians to coin the concept in this language) and one admitted term (the most suitable lay term used by consumers and patients to coin the concept in that language) for each participating language. In case the concept has a SNOMED-CT exact match, a “standardized term” is also given (literal translation of the fully specified SNOMED-CT name).

We made a deliberate choice not to represent hierarchical (broader- narrower) links or links between related concept into our system, nor any other attempt to self-generated ontological mapping. We decided to rely on the mappings to the international nomenclatures and classifications to explore the semantic relations between the concepts in our resource, and not to invest energy in the possibly redundant activity of creating a new ontology.

3.2. Creating an LMF based end-user lexicon

The second type of resource in the interface terminology is a series of unilingual end-user lexicons that must be linked to the multilingual reference terminology described above. These end-user lexicons (one for each language) can be linked to Natural Language Processing (NLP) applications, and can be oriented to patients and to professionals.

To represent the end-user lexicons, we used the ISO standard LMF[16], a model that provides a common standardized framework for NLP lexicons. This model deals with linguistic complexities, and, as TMF, uses the ISOcat source for the association of linguistic data categorie (e.g., *partOfSpeech*, *namedEntity*). Furthermore it can be linked to TMF and other concept based representation systems. Resources in LMF can also be linked to existing lexical NLP resources (such as WordNet). LMF meta-model contains a mechanism to deal with multiple senses (Francopoulo et al, 2007).

Lexical entries of the end-user lexicon are words (or phrases) that are selected from everyday interactions between doctors and patients (occurrences in medical records, in guidelines, in web consultations, etc.), based on frequency count and relevance. Various methodologies for human or automated term extraction can be used. For each of the selected words (or phrases), the possible senses (i.e. medical and non-medical ones) are clearly defined and entered as such in the LMF end-user lexicon of the originating language.

We propose a dual mechanism to link the first type of resource, the multilingual reference terminology and the end-user unilingual lexicons:

- Each concept in the reference terminology is linked to the sense part of a word in the lexicon. This mechanism preserves the conceptual integrity and is language independent.
- Each lexical representation, in a specific language, of a concept in the reference terminology, is linked to the corresponding lemma in a end-user lexicon of that language. This is language dependent.

In the unilingual end-user lexicons, links to synonyms for selected words (or phrases) in the selected sense can be provided. At this level, also subtle differences between related languages (e.g. Portuguese in Portugal and in Brazil, English in the UK and in the US, Dutch in the Netherlands and in Belgium) are addressed.

3.3. Publication in Linked Open Data

The TMF model of the reference terminology, described in Section 3.1, is implemented as an OWL-DL ontology, which is an efficient way of defining the components (represented as hierarchically organized classes) and the vocabulary (consisting of data and object properties that can be easily reused in other data sets) and ensuring the consistency of the data.

[16] http://www.lexicalmarkupframework.org/

Terminological entries are represented as classes, while data categories are represented as OWL object properties. Classes, whenever possible, are linked to similar concepts in other international recognized classifications, available on the web as Linked Data, using the *owl:equivalentClass* property (which, in this case, should be preferred to *owl:sameAs* to avoid undesired effects when using reasoners (Halpin et al., 2010). Because a one-to-one correspondence cannot always be found between all the classifications, in some cases entries can be grouped in more general categories with the *owl:unionOf* property.

The conversion of the TMF resource in OWL/RDF allows for the publication of the data as Linked Data[17] on the Semantic Web that are accessible via SPARQL queries. Linked Data principles encourage reuse, reduce redundancy, maximize its real and potential inter-connectedness, and finally enable network effects to add value to data (Bizer et al., 2009).

The end-user lexicons will be published according to the LEMON framework[18], which is a proposed model for modeling lexicons and machine-readable dictionaries based on LMF and, similarly to our TMF terminology, using Semantic Web technologies and ISOcat data categories (McCrae et al., 2012). The functionality to publish the content of the lexicons in Linked Open Data is an established part of the LEMON framework.

4 A use case for Heart Failure

To populate the structure of our interface terminology with a test sample, we have chosen to extract relevant concepts and words or phrases from a Belgian bilingual (Dutch and French) guideline on Heart Failure for general practitioners, published by the Scientific Associations of Primary Care Physicians (Van Royen et al., 2011). For this study we worked with the French version as a starting point. We describe a manual concept extraction and an automated term extraction, along with the procedures to populate both the TMF and LMF-LEMON resources.

[17] http://linkeddata.org/

[18] http://www.lemon-model.net/

4.1. Concept extraction to the TMF resource.

A general practitioner, expert in medical classifications, analyzed the French version of mentioned guideline, and after a careful reading and tagging, selected 168 concepts, relevant for the clinical domain of heart failure and pertaining to the reference world of general practice.

In a first step, all identified concepts were entered in the TMF resource, at the Term Entry section of the French language section.

A definition in French was given for each concept, together with a reference to the French guideline from which the concept was extracted. For each concept, a preferred term (representing the technical term used by physicians) and one or more admitted terms (representing the lay term used by patients for that concept) were chosen. For some technical terms the corresponding lay term was a simple description of the term itself (e.g. the French admitted term “eau dans le ventre” for the technical term “ascite”).

Then, in a second step the Terminological Entry level was addressed. The corresponding concepts were looked up in ICPC-2, the reference classification for general practice. Then, the cross mapping to ICD10 were sought after. Next a search in the SNOMED-CT web browser was made and finally the relevant corresponding definitions were extracted from the UMLS Metathesaurus, using a dedicated tool. In case a perfect match with a SNOMED-CT concept was possible, its Fully Specified Name was entered (as well as its French literal translation). In case the concept could not be matched in SNOMED-CT, a definition in English was sought (and the event recorded for notification to the international SNOMED-CT governance group IHTSDO).

At the end of this process, among the 168 concepts selected from the guideline, 153 were mapped to ICPC-2, 131 to ICD-10, 161 concepts to SNOMED-CT, and 116 to UMLS definitions.

Reasons for the inability to match with SNOMED-CT were the too broad and general nature of the selected concept, and mismatches between “world of references”. For instance, in the guideline, the concept “sexual problems” was repeatedly used at this broad level to convey information (and thus necessary in the communication). Some concepts (e.g. drug side effect “dry mouth”) were not within the scope of SNOMED-CT, as the category “Symptom and complaint" (very important for primary care) is alien to SNOMED-CT. Both UMLS and SNOMED-CT use semantic types to categorize concepts. However, the conceptual framework of these semantic types is different, which also leads to difficulties in finding an exact conceptual match for locally used medical terms in international nomenclatures. With regard to mapping to ICD-10, we noticed that 15 of the 168 concepts referred to aspects of functional status (e.g. nutritional status, exercise intolerance), which is difficult to represent in ICD. This provoked difficulties to map to this classification (oriented towards morbidity and mortality classification). Similarly, a number of terms related to medical procedures could not be mapped. Regarding ICPC-2, some of the selected concepts were mapped to more than one ICPC-2 rubric (e.g. "Alcoholic cardiomyopathy" mapped to "heart disease other" (K84) and "chronic alcohol abuse" (P15)) and in most cases they were mapped to broad rag bag ICPC-2 rubrics as in the case of the concept "maladie de Paget" mapped to the ICPC-2 rubric "T99 - Musculoskeletal other diseases" .

For each concept, also the Italian language section was populated with a translation of the preferred term and of the admitted term. This was performed with the help of Italian speaking domain experts and a terminologist, sustained by mapping to an existing Italian medical vocabulary oriented to healthcare consumers (the already mentioned ICMV). A similar work is to be made for English and Dutch language sections, and further expanded as described above.

4.2. Term extraction to the LMF resource.

The two versions (Dutch/French) of the heart failure guideline (30 pages of text) were submitted to the term extraction program TExSIS (Macken et al., 2013) , which provides tokenization, part-of-speech tagging, lemmatization, detection of phrases, named entity recognition, and bilingual sentence alignment. The term extraction resulted in an aligned bilingual glossary of 774 words and phrases.

After this automatic extraction, an exact matching to the French preferred terms of concepts in the TMF resource was performed. Surprisingly only 77 French preferred terms among 168 had a string match with the 774 word and phrases extracted by TExSIS. These 77 French terms were then entered in the French LMF - LEMON resource and linked to the identification number of the concepts in the TMF resource.

In addition, we considered all the 77 corresponding admitted (lay language[19]) terms from the TMF resource, resulting into an increase of entries in the LMF-LEMON resource to 138 lexical entries. For 16 admitted terms there was no difference with the French preferred term. Among the 138 lexical entries, 114 were phrases, and were then decomposed into single words, the total of entries to 298 in the LMF-LEMON resource. This rapid increase in the number of terms is however unlikely to be linear, since some words are part of several phrases (e.g. "insuffisance cardiaque" – "insuffisance rénale").

For each entry we added linguistic information, such as part-of-speech tag, lemma, spelling variants and inflected forms. The medical senses of the entries are linked to the reference terminology, and the other senses to DBpedia[20].

4.3. Publication of the TMF and LMF-LEMON resources in Linked Open Data

The tabular data representing the TMF resource on Heart Failure were automatically converted into 16.636 RDF triples, assigning a unique identifier to each element, and serializing it in RDF/XML. We published the resulting file (in French, English and Italian) on the Data Hub[21] under a Creative Commons License[22], so that it can be freely available for computer applications. The publication is provided using the Meriterm[23] server along with a SPARQL endpoint. By the combined use of URL rewriting techniques and SPARQL construct queries, triples are given a more convenient access.

---

[19] ''Lay'' language is the language used by patients to communicate between themselves or with physicians about their health problems, using for example ''heart attack'' instead of ''Myocardial Infarction'', as opposite to professional language which is the one used by doctors to communicate between themselves. For example CVA (Cardiovascular Accident) or TIA (Transient Ischemic attack) is useful between doctors but has to be "translated" in lay language for patient.

[20] http://fr.dbpedia.org/

[21] http://datahub.io/

[22] http://creativecommons.org/

[23] Medical Reference Interface Terminologies: http://meriterm.org

Also the LMF-LEMON resource is represented in RDF triples from the start, and hence automatically transferrable in Linked Open Data[24].
The following table show some statistics on the use case of Heart Failure.

	English	French	Italian
TMF Concepts		168	
Preferred Terms		168	168
Admitted Terms		168	168
Standardized Terms	161	161	161
UMLS definitions	116		
Links to ICD-10	131		
Links to ICPC-2	153		
TMF Triples	16,636		
LMF entries	298		
LMF Triples	3,400		

Table 1. Statistics on the Hearth Failure use case

5 Discussion and Conclusion

In this article, we presented the architectural structure of a multilingual medical interface terminology with an ambitious set of objectives. We presented a first attempt, within a small use case, to create a sophisticated, hybrid medical interface terminology, aiming at multilingual solutions, at semantic interoperability, and at open source availability as Linked Open Data in the Semantic Web. For resources as precious as language and international terminologies a proprietary approach would not be appropriate.
The value of our resource for medical communication is not tested yet in a clinical setting. Preliminary results indicate the need for a concerted quality control of the process of "words/concepts" selection. To improve results, a refinement of the mapping approach (manual, so time and cost consuming) is needed, trying to investigate semi-automatic approaches relying on Semantic Web technologies, as done in (Cardillo et al., 2012).

One big stimulus for our work is the increasing internationalization of health care (near border, cross border and intercontinental health care exchange including migrants health issues). In fact, in order to provide semantic interoperability between different healthcare information systems and between different health actors and patients, multilingual access to international terminologies is needed.

In conclusion, this terminology support system, relying on ISO standards and Semantic Web languages and tools, published as Linked Open Data supports: (i) the efficient use of existing medical terminologies and their legacy data in the activity of clinical encoding; (ii) links between professional language and lay language, and healthcare information system integration (e.g. between EHRs and PHRs); (iii) multilingualism in its core approach to semantic interoperability; (iv) information retrieval, and (v) creation of information through epidemiological research. Our interface terminology will enable physicians to find the right medical entry at the right moment at the point of care, and to integrate their data with standard classifications for their encoding, being consistent with the requirements of their Health Authority (e.g. the mandatory use in the General Practice of ICPC, or ICD as coding system for diagnoses or problem lists), and finally to exchange their data with other healthcare professionals and within various healthcare information systems in an interoperable way.

We are currently working on widening the coverage of our reference terminology selecting other use cases (e.g. contra-indications of medication) and to extend the range of participating languages.

We are also building a collaborative, web-based management platform for the extension and maintenance of the proposed Multilingual Medical Interface Terminology.

Acknowledgments

This study has been done under the MERITERM consortium, devoted to joined research activities on medical terminologies and classification systems. The authors wish to thank the consortium for the support.

References

Christian Bizer, Tom Heath, and Tim Berners-Lee. 2009. Linked data-the story so far, International Journal on Semantic Web and Information Systems, 5(3), 1-22.

Olivier Bodenreider. 2008. Comparing SNOMED CT and the NCI Thesaurus through Semantic Web Technologies. In proceedings of the 3rd International Conference on Knowledge Representation in Medicine (KRMED2008). R. Cornet, K.A. Spackman (Eds), vol. 410, pp. 37-43.

[24] The RDF/XML file with the lexicon and its metadata can be found online at the link: http://meriterm.org/heartfailure/lexicons.rdf/.

Elena Cardillo, Germano. Hernandez, and Olivier Bodenreider. 2012. Integrating consumer-oriented vocabularies with selected professional ones from the UMLS using Semantic Web Technologies. In M. Szomszor, P. Kostkova, O. Akan, P. Bellavista, J. Cao, D. Jiannong, F. Dressler, D. Ferrari, M. Gerla, H. Kobayashi, S. Palazzo, S. Sahni, X. Shen, M. Stan, J. Xiaohua, A. Zomaya, G. Coulson (eds.), Electronic Healthcare, Proceedings of e-Health2010, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 69, Part 4, pp. 54-61, Springer Berlin Heidelberg.

Elena Cardillo. 2011. A Lexi-ontological Resource for Consumer Healthcare: the Italian Consumer-oriented Vocabulary. PhD thesis, University of Trento, Fondazione Bruno Kessler. http://eprints-phd.biblio.unitn.it/570/.

Gil Francopoulo, Nuria Bel, Monte George, Nicoletta Calzolari, Monica Monachini, Mandy Pet, and Claudia Soria. 2007. Lexical Markup Framework: ISO standard for semantic information in NLP lexicons. In Proceedings of the Workshop on Lexical-Semantic and Ontological Resources of the GLDV Working Group on Lexicography at the Biennal Spring Conference at the GLDV, Tubingen, (13-14/04/2007).

Kin W. Fung and Olivier Bodenreider. 2005. Utilizing the UMLS for semantic mapping between terminologies. In Proceedings of AMIA Annual Symposium: 266–270.

Harry Halpin, Ivana Herman, and Patrick J. Hayes. 2010. When owl:sameAs isn't the Same: An Analysis of Identity Links on the Semantic Web. RDF Next Steps Workshop, Palo Alto, USA.

International Standard Organization. 1999. ISO 12620:1999 Computer applications in terminology – Data categories.

International Standard Organization. 2003. ISO 16642:2003 Computer applications in terminology - Terminological markup framework (TMF).

International Standard Organization. 2008. ISO 24613:2008 Language resource management — Lexical markup framework (LMF).

Lieve Macken, Els Lefever, and Veronique Hoste (2013). TExSIS: Bilingual terminology extraction from parallel corpora using chunk-based alignment. Terminology, Volume 19(1):1-30.

John McCrae, Guadalupe Aguado-de Cea, Paul Buitelaar, Philipp Cimiano, Thierry Declerck, Asunción Gómez-Pérez, Jorge Gracia, Laura Hollink, Elena Montiel-Ponsoda, Dennis Spohr, and Tobias Wunner (2012). Interchanging lexical resources on the semantic web. Language Resources and Evaluation, 46(4):701–719.

Majid Khayari, Stéphane Schneider, Isabelle Kramer, and Laurent Romary. 2006. Unification of multilingual scientific terminological resources using the ISO 16642 standard. The TermSciences initiative, LREC 2006, International Workshop Acquiring and representing multilingual, specialized lexicons: the case of biomedicine. http://hal.archives-ouvertes.fr/hal-00022424.

Alan. L. Rector and Sebastian Brandt. 2008. Why do it the hard way? The case for an expressive description logic for SNOMED. Journal of American Medical Informatics Association, 15(6):744–751.

Laurent Romary, Isabelle Kramer, Susanne Salmon-Alt, and Joseph Roumier. 2006. Gestion de données terminologiques: principes, modèles, methods. Terminologie et accès à l'information. Widad Mustafa El Hadi (Ed.), Hermes science publ, Paris, France.

S. Trent Rosenbloom, Randolph. A. Miller, Kevin B. Johnson, Peter I. Elkin, and Steven H. Brown. 2006. Interface terminologies: Facilitating direct entry of clinical data into electronic health record systems. J Am Med Inform Assoc, May-June, 13(3):277-288.

Joseph Roumier, Robert Vander Stichele, Laurent Romary, and Elena Cardillo. 2011. Approach to the Creation of a Multilingual, Medical Interface Terminology. Workshop Proceedings of the 9th International Conference on Terminology and Artificial Intelligence, 13-15.M.

Tania Tudorache, Sean M. Falconer, Csongor Nyulas, Natalya F. Noy, and Marc A. Musen. 2010. Will semantic web technologies work for the development of icd-11? In Proceedings of the International Semantic Web Conference (ISWC2010): 257–272.

Paul Van Royen, Pierre Chevalier, Gilles Dekeulenaer, Martine Goossens, Philip Koeck, Michel Vanhalewyn, and Paul Van den Heuvel. 2011. Recommandation de Bonne Pratique Insuffisance cardiaque. Domus Medica. SSMG. Belgium.

Qing Zeng and Tony Tse. 2006. Exploring and Developing Consumer Health Vocabularies. Journal of the American Medical Informatics Association,13:24–29.