Applied Ontology 7 (2012) 375–386
DOI 10.3233/AO-2012-0119
IOS Press

Guest Editorial

Ontologies and terminologies: Continuum or dichotomy?

Natalia Grabar [a,*], Thierry Hamon [b] and Olivier Bodenreider [c]
[a] *CNRS UMR, Université Lille 1&3, Lille, France.*
E-mail: natalia.grabar@univ-lille3.fr
[b] *LIM&BIO UFR SMBH, Université Paris 13, Paris, France*
E-mail: thierry.hamon@univ-paris13.fr
[c] *National Library of Medicine, Bethesda, MD, USA*
E-mail: obodenreider@nlm.nih.gov

Abstract. Since there is a great confusion between the ontologies and other semantic resources, the purpose of this special issue is to address the question on "Ontologies and terminologies: Continuum or dichotomy". We have selected five articles which study the differences and similarities between these semantic resources. More particularly, the articles are dedicated to the differences existing at the level of terms and of relations, the use of the ontologies on corpora and the dynamic and static representation of the knowledge.

Keywords: Terminology, ontology, distinction, entities, relations, knowledge, continuum

1. Introduction

Storage and management of our knowledge of the world or of a specific domain is a very old endeavour started by Aristotle, pursued by the Encyclopedists, Wüster and the Vienna School, and more recently reinvigorated with the development of Artificial Intelligence. Various types of artifacts have been created to capture this knowledge, including lexica, dictionaries, universal and domain ontologies, taxonomies and terminologies. The recent development of and increasing interest for ontologies has somewhat blurred the distinction among these different types of artifacts and especially between ontologies and other semantic resources. In fact, as we will illustrate, the term ontology is used so liberally that it is not always clear what the differences among the various types of semantic resources are. Research on ontologies has attempted to elicit the distinctions among these resources (Guarino et al., 2009; Schulz & Stenzhorn, 2007; van Rees, 2003).

- A terminology is usually defined as a set of terms, which represent the system of concepts for an area or for an application. Terms are linguistic entities and linguistic information may be associated

*Corresponding author: Natalia Grabar, CNRS UMR 8163 STL, Université Lille 1&3, BP60149, F-59653 Villeneuve d'Ascq Cedex, France. E-mail: natalia.grabar@univ-lille3.fr.

with them. Term organization is usually not constrained by any formal logic, which may lead to problems like cyclicity and redundancy within a terminology.
- An ontology also describes a system of concepts and its associated properties for a specific area. However, ontologies are built upon formal specification and constraints. They are often intended to support computer applications.

Nevertheless, because of the proximity in their semantic content and application contexts, terminologies and ontologies may be better thought of as part of a continuum rather than completely distinct types of artifacts. The following examples highlight the lack of precise distinction among semantic resources in the scientific literature. We show that this confusion applies to resources for both general language and specialized domains (biomedicine and agriculture).

- The general language resource *WordNet* (Fellbaum, 1998):
  - is very often considered a lexical database (Bobillo et al., 2012; Burgun & Bodenreider, 2001; Fragos et al., 2008; Gangemi et al., 2003a, 2003b; Hearst, 1998; Lüngen & Storrer, 2007; Miller, 1995; Miller & Hristea, 2006; Niles & Pease, 2003; Wong, 2004),
  - but also a (linguistic) ontology (Alfonseca & Manandhar, 2002; Amshakala & Nedunchezhian, 2011; Asanoma, 2001; Boudhh & Bhattacharyya, 2010; Buscaldi et al., 2006; Djuana et al., 2011; Hung & Wermter, 2004; Ingvaldsen & Veres, 2004; Moravec et al., 2004; O'Hara et al., 1998; Schweighofer & Liebwald, 2007; Suchanek. et al., 2008),
  - or a folksonomy (Djuana et al., 2011).
  - In some articles it is also associated with several of them (Amshakala & Nedunchezhian, 2011; Schweighofer & Liebwald, 2007).
- *Foundational Model of Anatomy* (FMA) (Rosse & Mejino, 2003):
  - is mainly presented as ontology (Aleksovski et al., 2006; Burgun, 2006; Jimenez-Ruiz et al., 2012; Mizoguchi et al., 2009; Rickard et al., 2004; Smith et al., 2007; Zhang & Bodenreider, 2003),
  - but is considered a terminology by researchers interested primarily in text processing (Buitelaar et al., 2008; Grosjean et al., 2011).
- *Medical Subject Headings* (MeSH) (MeSH, 1998):
  - is usually described as a terminology or thesaurus (Claveau and Kijak, 2011; Grabar & Zweigenbaum, 2002; Hardcastle & Hallett, 2007; Merabti et al., 2012; Mottaz et al., 2008; Névéol et al., 2004),
  - but is also sometimes called an ontology (Abasolo & Gomez, 2000; Bloehdorn & Hotho, 2004; Ginter et al., 2004; Jiménez-Ruiz & Grau, 2011; Mörchen et al., 2008; Yoo & Hu, 2006).
  - Moreover, in several publications *MeSH* is called ontology but is described as a thesaurus or a controlled vocabulary (Aleksovski et al., 2006; Elberrichi et al., 2012; Hliaoutakis et al., 2006; Petrakis et al., 1998; Rak et al., 2008).
- *AgroVoc* resource (Food and Agriculture Organization of the United Nations, 1995):
  - is mainly described as a thesaurus or as a structured controlled vocabulary (Autayeu, 2011; Daille, 2000; Gangemi et al., 2002; Jacquemin et al., 1997; Kawtrakul et al., 2005; Lauser et al., 2008; Liang et al., 2006; Medelyan & Witten, 2008),

  - but is also called an ontology by some authors (Bloehdorn & Hotho, 2004; Gracia et al., 2007; Patwar et al., 2009), who do not even mention that it has indeed been reengineered into an ontology (Soergel et al., 2004).

- *NCI thesaurus* (Golbeck et al., 2003):
  - is defined by some authors as a terminology or thesaurus (Bodenreider, 2008; de Coronado et al., 2004),
  - but also as an ontology (Beltrán et al., 2010; Jiménez-Ruiz et al., 2008; Jiménez-Ruiz & Grau, 2011; Lambrix et al., 2009; Marquet et al., 2007; Ongenae et al., 2010),
  - or as both a thesaurus and an ontology (Goncalves et al., 2011; Fragoso et al., 2004, Kementsietsidis et al., 2008; Noy et al., 2008).

- *SNOMED CT* (IHTSDO, 2008):
  - is called a terminology (Ceusters et al., 2004; Lee et al., 2010; Patrick et al., 2007; Qamar & Rector, 2007; Schulz et al., 2007; Spackman et al., 1997),
  - an ontology (Jiménez-Ruiz & Grau, 2011; Halland et al., 2011; Kazakov, 2009; Milian et al., 2010; Patrick, 2006),
  - or both (Haase & Lutz, 2008; Nguyen et al., 2009; Seyed et al., 2012; Simon et al., 2004).

- *Unified Medical Language System* (UMLS) (Lindberg et al., 1993):
  - is mostly characterized as a meta-thesaurus or domain-specific terminological system (Achour et al., 1999; Besana et al., 2010; Erdogan et al., 2010; Hettne et al., 2010; Kumar et al., 2003; McInnes et al., 2009; Patel & Cimino, 2006; Schulze-Kremer et al., 2004; Wu et al., 2012).
  - However, in several articles it is referred to as an ontology (Cešpivová et al., 2004; Jonquet et al., 2009; Lasbleiz et al., 2005; Spasic et al., 2005; Vintar et al., 2003),
  - or both (Merrill et al., 2008; Pérez-Rey et al., 2006; Pustejovsky et al., 2002).

These few examples clearly illustrate the confusion between ontologies and other types of semantic resources in the scientific literature. Another interesting observation is that the criteria for labeling a given semantic resource an ontology, a terminology or something else are rarely explicit. For example, the mere presence of a hierarchical structure is sometimes sufficient for calling a resource an ontology, even if it is not a discriminating feature of ontologies. Besides, several researchers have pointed out the ontological limitations of *WordNet* (Gangemi et al., 2003b; Masolo et al., 2003), *MeSH* (Kibbe & Schriml, 2008), *NCI thesaurus* (Ceusters et al., 2005), *SNOMED CT* (Ceusters et al., 2004; Heja et al., 2008; Schulz & Cornet, 2009; Schulz et al., 2007; Spackman & Reynoso, 2004) and *UMLS* (Erdogan et al., 2010; Kumar & Smith, 2003a, 2003b; Kumar et al., 2003; Pisanelli et al., 1998, 1999; Schulze-Kremer et al., 2004). In contrast, the *Foundational Model of Anatomy* resource seems to satisfy ontological requirements, and *Agrovoc*, which has been reengineered into an ontology (Soergel et al., 2004), now exists in two flavors (thesaurus and ontology). Actually, it has been suggested that semantic resources, such as thesauri and terminologies, can form the basis for developing ontologies (Gangemi et al., 2002; Kumar & Smith, 2005). But the process of building an ontology from such resources requires their formalization and structure modification in order to prevent semantic inconsistencies (Gangemi et al., 2002; Smith et al., 2005; van Assem et al., 2006).

2. Objectives of this special issue

In this special issue of *Applied Ontology*, we propose to address various issues related to differences and similarities between ontologies and other semantic resources in order to reduce the confusion between these types of artifacts. The next section introduces the contributions we have collected for this issue.

The selection was very competitive and we could accept only five articles for publication. For this reason, several aspects are not addressed in this special issue, such as:

- Automatic and semi-automatic acquisition of semantic resources (Aussenac-Gilles, 2005; Drouin, 2002; Maedche & Staab, 2000).
- Adaptation and modularity of semantic resources (Jurisica et al., 2004; Stenzhorn et al., 2008; Yu, 2006).
- Building ontologies from existing semantic resources, such as thesauri or terminologies (Gangemi et al., 2002; Kumar & Smith, 2005; Smith et al., 2005; van Assem et al., 2006).
- Reengineering of terminologies and thesauri into ontologies (Soergel et al., 2004).
- Pre-coordination and post-coordination within semantic resources (Rector et al., 2009; Spackman & Campbell, 1998).
- Use of verbs *vs.* nouns in knowledge representation (Lerat, 2002; L'Homme, 1998, 2012; Lorente, 2002; Pimentel, 2011).
- Impact of formal, structural and content differences between different kinds of semantic resources on their use, as well as on the results provided by automatic systems.
- Multilingual aspect of semantic resources and their localization.
- . . . .

3. The contributions collected in this issue

We have received a total of 17 submissions. A two-step selection and a rigorous reviewing process by at least three members of the scientific committee led to the selection of the five submissions published in this issue (30% acceptance rate). The main findings of these five articles are summarized here. The interested reader is referred to the full articles for details.

3.1. Terms as labels for concepts, terms as lexical units: A comparative analysis in ontologies and specialized dictionaries

The article by Marie-Claude L'Homme and Gabriel Bernier-Colborne analyses the differences observed at the level of terms in two types of semantic resources: specialized dictionaries (terminologies) and ontologies. The objective of this work is to understand what a term is and what role it plays in these two kinds of semantic resources. More specifically, this work shows that:

- In ontologies, terms may be viewed as devices to access individual concepts. Since other non-linguistic devices (such as numerical identifiers) can have the same function, this role is secondary, but important, as terms also convey a cognitive dimension to ontologies. Terms have a dual status: they can be viewed as a means to express concepts or considered independent from the concepts.

- In specialized vocabularies, terms appear as headwords and necessarily correspond to linguistic expressions that occur in specialized texts or that are used by experts. Since the linguistic expressions cannot be created artificially, the terms are real expressions. Specialized dictionaries take into account both cognitive and linguistic dimensions of terms, albeit in varying proportions.

The authors conclude that ontologies and specialized dictionaries deal with different objects and that their objectives are different as well. Ontologies aim to organize canonical knowledge, which sometimes necessitates the use of abstract categories in order to better organize things and to facilitate their use by automatic tools. In contrast, dictionaries aim to describe terms that are not very well organized, exactly like the natural language to which they belong.

3.2. Relationships and relata in ontologies and thesauri: *Differences and similarities*

The article by Daniel Kless, Simon Milton and Edmund Kazmierczak also proposes a comparison between thesauri and ontologies, but at the level of the relations. This study analyzes the scientific and standardisation literature and some existing thesauri and ontologies. Several differences are pointed out:

- The relations are considered as properties in ontologies, while they play a linking role in thesauri. Hence, ontology relationships express necessary membership conditions for their relata.
- The thesaurus hierarchical relations may correspond to different kinds of relationships compared to the relations found in ontologies (*part-of*, *is-a* or *instance*).
- The relata of the thesauri relationships may be heterogeneous, while they are homogeneous in ontologies (universals, individuals, pure or mixed collections).
- The thesaurus hierarchical relationships are generally transitive and expressed bidirectionally, and the use of different kinds of relationships to create hierarchies may break the transitive property of hierarchical relations. In contrast, in ontology, the relationships do not imply the existence of an inverse relationships between two relata.

The authors conclude that thesauri and ontologies apparently have similar structures, but with different models. Some of these differences are at the level of the relationships.

3.3. Corpus-based terminological evaluation of ontologies

The article by Marco Rospocher, Sara Tonelli, Luciano Serafini and Emanuele Pianta discusses the evaluation of ontologies and of their relevance to corpora. This article highlights that the ontological level (logical organization) and its terminological content (term labels) are closely related. In text-based applications and tasks, the terminological level plays an important role, as has been stressed in the first article. Moreover, the authors propose a framework for the evaluation of the content coverage of an ontology through a corpus. They rely on existing tools (KX combined with a word sense disambiguation library) and resources (WordNet), and they also adapt the weighting scheme (recall and F-measure). The proposed framework and methods are tested with several ontologies and corpora.

3.4. The social life of categories: *An empirical study of term categorization*

The article by John Lamp and Simon Milton analyzes the categorization of terms according to the expertise level of participants. The study uses Boisot's model for the representation of knowledge discovery and appropriation, i.e. how the knowledge passes from a chaotic to an ordered state. The authors show

that there is a relation between the expertise level of users and Boisot's model: a wide range of views is held by participants about the nature and structure of things within a domain. For instance, when asked to categorize terms, the participants demonstrate a diversity of views (which terms belong to which categories? what is the number of relevant categories for each term?). Very interestingly, this experiment also shows that learning is a dynamic process: things move in Boisot's model as the participants' knowledge becomes more complete. With more advanced knowledge, the level of consensus in term categorization becomes higher for various features (number of terms per category, number of categories, hierarchy, proximity between terms, distribution of categories within semantic structures. . .). Finally, the authors observe that in ontologies and terminologies the categorization result is a static representation, whereas both the domain and our knowledge are dynamic and continue to evolve.

3.5. Beyond terminologies: *Using psychometrics to validate shared ontologies*

The article by Dirk van der Linden, Stijn Hoppenbrouwers, Alina Lartseva and Wolfgang Molnar investigates how various users or modelers interpret the meta-conceptual constructs (categories) in ontology modeling languages. The topic of this article is related to the previous article. More specifically, the authors show that, contrary to what is commonly accepted, there is no a priori consensus about the semantics, but that there is rather a specialization of discourse communities and of their individual members. The authors propose a method for measuring the variability of the categorization. They conclude that there is an explicit tendency towards the personal ontologies which account for interpersonal semantic differences and that the creation of shared ontologies may be a difficult objective to achieve.

Acknowledgements

Our thanks go to all the authors who committed their time and submitted their work to this special issue. We would like also to thank the members of the scientific committee for the reviews they provided for this special issue: Jiye Ai, Nathalie Aussenac, Paul Buitelaar Sylvie Després, Christiane D. Fellbaum, Anand Kumar, Nathalie Hernandez, Marie-Claude L'Homme, Véronique Malaisé, Fleur Mougin, Aurélie Névéol, Alessandro Oltramari, Chantal Reynaud, Stefan Schulz, Dagobert Sörgel, Rita Temmerman, Maria Teresa Pazienza, Hanne E. Thomsen, Susan Thomas, Anna Tordai, Karin Verspoor. Above all, we are grateful to Roberta Ferrario, Nicola Guarino and Mark Musen, for their patience and help in the organization of this issue.

References

Abasolo, J.M. & Gomez, M. (2000). MELISA: an ontology-based agent for information retrieval in medicine. In *Proceedings of the 1st International Workshop on the Semantic Web* (pp. 73–82).

Achour, S., Dojat, M., Brethon, J.-M., Blain, G. & Lepage, E. (1999). The use of the UMLS knowledge sources for the design of a domain specific ontology: a practical experience in blood transfusion. In *Proceedings of the Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making* (pp. 249–253). London: Springer-Verlag.

Aleksovski, Z., ten Kate, W. & van Harmelen, F. (2006). Ontology matching using comprehensive ontology as background knowledge. In *Proceedings of the International Workshop on Ontology Matching at ISWC 2006* (pp. 13–24).

Alfonseca, E. & Manandhar, S. (2002). Extending a lexical ontology by a combination of distributional semantics signatures. In *Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web* (pp. 1–7). London: Springer-Verlag.

Amshakala, K. & Nedunchezhian, R. (2011). WordNet ontology based query reformulation and optimization using disjunctive clause elimination. *International Journal of Database Management Systems*, *3*(4), 55–63.

Asanoma, N. (2001). Alignment of ontologies: WordNet and Goi-Taikei. In *Proceedings of the NAACL Workshop, "WordNet and Other Lexical Resources: Applications, Extensions and Customizations"*, Pittsburg, PA, USA.

Aussenac-Gilles, N. (2005). Supervised text analysis for ontology and terminology engineering. In *Proceedings of the Dagstuhl Seminar on Machine Learning for the Semantic Web* (pp. 35–46).

Autayeu, A. (2011). Large scale semantic matching: Agrovoc vs CABI. In *Proceedings of the IJCAI Workshop on Discovering Meaning On the Go in Large Heterogeneous Data 2011 (LHD)*, Barcelona.

Beltrán, A.G., Tagger, B. & Finkelstein, A. (2010). Ontology-based queries over cancer data. In *Proceedings of the Workshop on Semantic Web Applications and Tools for Life Sciences*, Berlin.

Besana, P., Cuggia, M., Zekri, O., Bourde, A. & Burgun, A. (2010). From French EHR to NCI ontology via UMLS. In P. Shvaiko, J. Euzenat, F. Giunchiglia, H. Stuckenschmidt, M. Mao and I.F. Cruz (Eds.), *Proceedings of the 5th International Workshop on Ontology Matching (OM-2010)*, Shanghai.

Bloehdorn, S. & Hotho, A. (2004). Boosting for text classification with semantic features. In *Proceedings of the 6th International Conference on Knowledge Discovery on the Web: Advances in Web Mining and Web Usage Analysis*, Seattle, WA, USA (pp. 149–166).

Bobillo, F., Gómez-Romero, J. & León-Araúz, P. (2012). Fuzzy ontologies for specialized knowledge representation in WordNet. In S.G.B. Bouchon-Meunier, G.C.M. Fedrizzi, B. Matarazzo and R. Yager (Eds.), *Proceedings of the 14th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'2012), Part I*. Communications in Computer and Information Science (Vol. 297, pp. 430–439). Berlin/Heidelberg: Springer-Verlag.

Bodenreider, O. (2008). Comparing SNOMED CT and the NCI thesaurus through semantic web technologies. In *Proceedings of the 3rd International Conference on Knowledge Representation in Medicine*, Phoenix, AZ, USA.

Boudhh, S. & Bhattacharyya, P. (2010). Unification of universal word dictionaries using WordNet ontology and similarity measures. In *Proceedings of the 5th International Conference on Global WordNet (GWC2010)*, Mumbai.

Buitelaar, P., Wennerberg, P.O. & Zillner, S. (2008). Statistical term profiling for query pattern mining. In *Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing* (pp. 114–115). Columbus, OH, USA: Association for Computational Linguistics.

Burgun, A. (2006). Desiderata for domain reference ontologies in biomedicine. *Journal of Biomedical Informatics*, *39*(3), 307–313.

Burgun, A. & Bodenreider, O. (2001). Comparing terms, concepts and semantic classes in WordNet and the Unified Medical Language System. In *Proceedings of the NAACL Workshop, "WordNet and Other Lexical Resources: Applications, Extensions and Customizations"* (pp. 77–82).

Buscaldi, D., Rosso, P. & Arnal, E.S. (2006). Using the WordNet ontology in the GeoCLEF geographical information retrieval task. In *Proceedings of the 6th International Conference on Cross-Language Evaluation Forum: Accessing Multilingual Information Repositories* (pp. 939–946). Berlin/Heidelberg: Springer-Verlag.

Ceusters, W., Smith, B. & Goldberg, L. (2005). A terminological and ontological analysis of the NCI thesaurus. *Methods of Information in Medicine*, *44*, 498–507.

Ceusters, W., Smith, B., Kumar, A. & Dhaen, C. (2004). Ontology-based error detection in SNOMED-CT. In *Proceedings of Medinfo 2004* (pp. 482–486).

Cešpivová, H., Rauch, J., Svátek, V., Kejkula, M. & Tomecková, M. (2004). Roles of medical ontology in association mining CRISP-DM cycle. In *Proceedings of the ECML/PKDD04 Workshop on Knowledge Discovery and Ontologies (KDO'04)*, Pisa.

Claveau, V. & Kijak, E. (2011). Morphological analysis of biomedical terminology with analogy-based alignment. In *Proceedings of Recent Advances in Natural Language Processing*, Hissar, Bulgaria (pp. 347–354).

Daille, B. (2000). Morphological rule induction for terminology acquisition. In *Proceedings of the 18th International Conference on Computational Linguistics (COLING'2000)* (pp. 215–221).

de Coronado, S., Haber, M.W., Sioutos, N., Tuttle, M.S. & Wright, L.W. (2004). NCI thesaurus: using science-based terminology to integrate cancer research results. In *Proceedings of MEDINFO'2004* (pp. 33–37).

Djuana, E., Xu, Y. & Li, Y. (2011). Constructing tag ontology from folksonomy based on WordNet. In P. Kommers, J. Zhang, T. Issa & P. Isaías (Eds.), *Proceedings of the IADIS International Conference on Internet Technologies and Society 2011*, Shanghai.

Drouin, P. (2002). Acquisition automatique des termes: l'utilisation des pivotslexicaux spécialisés. PhD thesis, Université de Montréal.

Elberrichi, Z., Taibi, M. & Belaggoun, A. (2012). Multilingual medical documents classification based on mesh domain ontology. *International Journal of Computer Science Issues*, *9*(2), 150.

Erdogan, H., Erdem, E. & Bodenreider, O. (2010). Exploiting UMLS semantics for checking semantic consistency among UMLS concepts. In *Proceedings of the 13th International Congress on Medical Informatics (MedInfo'10)* (pp. 749–753).

Fellbaum, C. (Ed.) (1998). *WordNet: An Electronic Lexical Database*. Cambridge, MA, USA: MIT Press.

Food and Agriculture Organization of the United Nations (1995). Agrovoc multilingual agricultural thesaurus.

Fragos, K., Maistros, Y., Skourlas, C. & Skourlas, C. (2008). Using wordnet lexical database and internet to disambiguate word senses. In *Proceedings of the 9th Panhellenic Conference in Informatics*, Thessaloniki, Greece.

Fragoso, G., de Coronado, S., Margaret, H., Frank, H. & Larry, W. (2004). Overview and utilization of the NCI thesaurus. *Comparative and Functional Genomics*, *5*(8), 648–654.

Gangemi, A., Fisseha, F., Pettman, I., Pisanelli, D.M., Taconet, M. & Keizer, J. (2002). A formal ontological framework for semantic interoperability in the fishery domain. In *Proceedings of International Semantic Web Conference (ISCW'2002)*, Sardinia, Italy.

Gangemi, A., Guarino, N., Masolo, C., Oltramari, A. & Oltramari, R. (2003a). Sweetening WORDNET with DOLCE. *AI Magazine*, *24*(3), 13–24.

Gangemi, A., Navigli, R. & Velardi, P. (2003b). The OntoWordNet project: extension and axiomatization of conceptual relations in WordNet. In R. Meersman, Z. Tari and D.C. Schmidt (Eds.), *On the Move to Meaningful Internet Systems 2003: CoopIS, DOA & ODBASE – OTM Confederated International Conferences, CoopIS, DOA & ODBASE 2003*, Catania, Sicily, Italy. Lecture Notes in Computer Science (Vol. 2888, pp. 820–838). Berlin/Heidelberg: Springer.

Ginter, F., Pyysalo, S., Boberg, J., Järvinen, J. & Salakoski, T. (2004). Ontology-based feature transformations: a data-driven approach. In *Advances in Natural Language Processing (4th International Conference in NLP)*. Lecture Notes in Computer Science (Vol. 3230, pp. 280–291). Berlin/Heidelberg: Springer.

Golbeck, J., Fragoso, G., Hartel, F., Hendler, J., Oberthaler, J. & Parsia, B. (2003). The national cancer institute's thesaurus and ontology. *Journal of Web Semantics*, *1*, 75–80.

Goncalves, R.S., Parsia, B. & Parsia, U. (2011). Analysing the evolution of the NCI thesaurus. In *Proceedings of the 24th International Symposium on Computer-Based Medical Systems* (pp. 1–6). Washington, DC, USA: IEEE Computer Society.

Grabar, N. & Zweigenbaum, P. (2002). Lexically-based terminology structuring: some inherent limits. In *Proceedings of Computerm'2002 (2nd Workshop on Computational Terminology)*, Taiwan (pp. 1–7).

Gracia, J., Lopez, V., d'Aquin, M., Sabou, M., Motta, E. & Mena, E. (2007). Solving semantic ambiguity to improve semantic web based ontology matching. In *Proceedings of the 2nd International Workshop on Ontology Matching (OM-2007)*, Busan, South Korea.

Grosjean, J., Merabti, T., Dahamna, B., Kergourlay, I., Thirion, B., Soualmia, L.F. & Darmoni, S.J. (2011). Health multi-terminology portal: a semantics added-value for patient safety. In *Patient Safety Informatics – Adverse Drug Events, Human Factors and IT Tools for Patient Medication Safety*. Studies in Health Technology and Informatics (Vol. 166, pp. 129–138).

Guarino, N., Oberle, D. & Staab, S. (2009). What is an Ontology? In S. Staab and R. Studer (Eds.), *Handbook on Ontologies*, International Handbooks on Information Systems (pp. 1–17). Berlin/Heidelberg: Springer-Verlag.

Haase, C. & Lutz, C. (2008). Complexity of subsumption in the $\mathcal{EL}$ family of description logics: acyclic and cyclic TBoxes. In *Proceedings of ECAI'2008* (pp. 25–29).

Halland, K., Britz, K. & Gerber, A. (2011). Investigations into the use of SNOMED CT to enhance an OpenMRS health information system. *South African Computer Journal*, *47*, 33–45.

Hardcastle, D. & Hallett, C. (2007). Exploring the use of NLP in the disclosure of electronic patient records. In *Biological, Translational & Clinical Language Processing*, Prague (pp. 161–162). Stroudsburg, PA, USA: Association for Computational Linguistics.

Hearst, M.A. (1998). Automated discovery of WordNet relations. In C. Fellbaum (Ed.), *WordNet: An Electronic Lexical Database*. Language, Speech and Communication (Chapter 5, pp. 131–151). Cambridge, MA, USA: MIT Press.

Heja, G., Surjan, G. & Varga, P. (2008). Ontological analysis of SNOMED CT. *BMC Medical Informatics and Decision Making*, *8*(Suppl. 1), S8.

Hettne, K., van Mulligen, E., Schuemie, M., Schijvenaars, B. & Kors, J. (2010). Rewriting and suppressing UMLS terms for improved biomedical term identification. *Journal of Biomedical Semantics*, *1*(1), 5.

Hliaoutakis, A., Varelas, G., Voutsakis, E., Petrakis, E.G.M. & Milios, E. (2006). Information retrieval by semantic similarity. *International Journal on Semantic Web and Information Systems*, *3*(3), 55–73.

Hung, C. & Wermter, S. (2004). Neural network based document clustering using wordnet ontologies. *International Journal of Hybrid Intelligent Systems*, *1*(3,4), 127–142.

IHTSDO (2008). SNOMED clinical terms – technical reference guide. International Health Terminology Standards Development Organisation. Available at: http://www.ihtsdo.org/snomed-ct.

Ingvaldsen, J.E. & Veres, C. (2004). Using the WordNet ontology for interpreting medical records. In *Proceedings of the CAiSE Workshops (3)* (pp. 355–358).

Jacquemin, C., Klavans, J.L. & Tzoukermann, E. (1997). Expansion of multi-word terms for indexing and retrieval using morphology and syntax. In *Proceedings of the ACL'97/EACL'97*, Barcelona (pp. 24–31).

Jiménez-Ruiz, E. & Grau, B.C. (2011). Towards more challenging problems for ontology matching tools. In *Proceedings of the 6th International Workshop on Ontology Matching*, Bonn, Germany.

Jiménez-Ruiz, E., Grau, B.C., Sattler, U., Schneider, T. & Llavori, R.B. (2008). Safe and economic re-use of ontologies: a logic-based methodology and tool support. In *Proceedings of the 21st International Workshop on Description Logics (DL2008)*, Dresden, Germany.

Jimenez-Ruiz, E., Grau, B.C., Zhou, Y. & Horrocks, I. (2012). Large-scale interactive ontology matching: algorithms and implementation. In *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI)*, Montpellier, France.

Jonquet, C., Shah, N.H., Youn, C.H., Callendar, C., Storey, M.-A. & Musen, M.A. (2009). NCBO annotator: semantic annotation of biomedical data. In *Proceedings of the 8th International Semantic Web Conference, Poster and Demonstration Session*, Washington DC, USA.

Jurisica, I., Mylopoulos, J. & Yu, E. (2004). Ontologies for knowledge management: an information systems perspective. *Knowledge and Information Systems*, *6*(4), 380–401.

Kawtrakul, A., Imsombut, A., Thunkijjanukit, A., Soergel, D., Liang, A., Sini, M., Johannsen, G. & Keizer, J. (2005). Automatic term relationship cleaning and refinement for AGROVOC. In *Workshop on the 6th Agricultural Ontology Service*, Vila Real, Portugal.

Kazakov, Y. (2009). Consequence-driven reasoning for horn SHIQ ontologies. In *Proceedings of the 22nd International Workshop on Description Logics*, Oxford, UK.

Kementsietsidis, A., Lim, L. & Wang, M. (2008). Supporting ontology-based keyword search over medical databases. In *Proceedings of the AMIA Annual Symposium* (pp. 409–413).

Kibbe, W. & Schriml, L. (2008). Towards an open source disease ontology. In *Proceedings of the Disease Ontologies and Information Workshop*. Hinxton, Cambridge, UK: WTCC.

Kumar, A., Ciccarese, P., Quaglini, S., Stefanelli, M., Caffi, E. & Boiocchi, L. (2003). Relating UMLS semantic types and task-based ontology to computer-interpretable clinical practice guidelines. In I. Press (Ed.), *Proceedings of MIE 2003* (pp. 469–474).

Kumar, A. & Smith, B. (2003a). Ontology for task-based clinical guidelines and the theory of granular partitions. In M. Dojat, T.E. Keravnou and P. Barahona (Eds.), *Proceedings of 9th Conference on Artificial Intelligence in Medicine Europe*. Lecture Notes in Computer Science (Vol. 2780, pp. 71–75). Berlin: Springer-Verlag.

Kumar, A. & Smith, B. (2003b). The unified medical language system and the gene ontology: some critical reflections. In A. Günter, R. Kruse and B. Neumann (Eds.), *Proceedings of KI2003 (Advances in AI)*. Lecture Notes in Computer Science (Vol. 2821, pp. 135–148). Berlin: Springer-Verlag.

Kumar, A. & Smith, B. (2005). Oncology ontology in the NCI thesaurus. In *Proceedings of the 10th Conference on Artificial Intelligence in Medicine* (pp. 213–220). Berlin/Heidelberg: Springer-Verlag.

Lambrix, P., Liu, Q. & Tan, H. (2009). A system for repairing missing is-a structure in ontologies. In *Proceedings of the Workshop on Semantic Web Applications and Tools for Life Sciences*, Amsterdam.

Lasbleiz, J., Bertaud-Gounot, V., Mougin, F., Burgun, A. & Duvauferrier, R. (2005). Ontologies and information models: comparing the UMLS and the DICOM content mapping resource. *Acta Informatica Medica*, *13*(4), 198–200.

Lauser, B., Johannsen, G., Caracciolo, C., van Hage, W.R., Keizer, J. & Mayr, P. (2008). Comparing human and automatic thesaurus mapping approaches in the agricultural domain. In *Proceedings of the 2008 International Conference on Dublin Core and Metadata Applications*, Berlin, Germany (pp. 43–53). Dublin Core Metadata Initiative.

Lee, D., Lau, F. & Quan, H. (2010). A method for encoding clinical datasets with SNOMED CT. *BMC Medical Informatics and Decision Making*, *10*, 53.

Lerat, P. (2002). Qu'est-ce que le verbe spécialisé? le cas du droit. *Cahiers de lexicologie*, *80*, 201–211.

L'Homme, M.-C. (1998). Le statut du verbe en langue de spécialité et sa description lexicographique. *Cahiers de lexicologie*, *73*(2), 61–84.

L'Homme, M.-C. (2012). Le verbe terminologique: un portrait des travaux récents. In *CMLF'2012* (pp. 93–107).

Liang, A.C., Lauser, B., Sini, M., Keizer, J. & Katz, S. (2006). From AGROVOC to the agricultural ontology service/concept server: an OWL model for creating ontologies in the agricultural domain. In *Proceedings of the 2006 International Conference on Dublin Core and Metadata Applications: Metadata for Knowledge and Learning*, Manzanillo, Colima, Mexico (pp. 76–88). Dublin Core Metadata Initiative.

Lindberg, D., Humphreys, B. & McCray, A. (1993). The unified medical language system. *Methods of Information in Medicine*, *32*(4), 281–291.

Lorente, M. (2002). Verbos y discurso especializado. *Estudios de linguistica espanola*, *16*. Available at: http://elies.rediris.es/elies16/Lorente.html.

Lüngen, H. & Storrer, A. (2007). Domain ontologies and WordNets in OWL: modelling options. *LDV Forum*, *22*(2), 1–17.

Maedche, A. & Staab, S. (2000). Mining ontologies from text. In *EKAW* (pp. 189–202). Heidelberg: Springer-Verlag.

Marquet, G., Dameron, O., Saikali, S., Mosser, J. & Burgun, A. (2007). Grading glioma tumors using OWL-DL and NCI thesaurus. In *Proceedings of the American Medical Informatics Association Conference* (pp. 508–512).

Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A. & Schneider, L. (2003). The wonderweb library of foundational ontologies and the DOLCE ontology. Wonderweb deliverable D17, Intermediate report, version 2.1.

McInnes, B.T., Pedersen, T. & Pakhomov, S.V. (2009). UMLS-interface and UMLS-similarity: open source software for measuring paths and semantic similarity. In *Proceedings of the AMIA Annual Symposium* (pp. 431–435).

Medelyan, O. & Witten, I.H. (2008). Domain-independent automatic keyphrase indexing with small training sets. *Journal of the American Society for Information Science and Technology*, *59*(7), 1026–1040.

Merabti, T., Soualmia, L.F., Grosjean, J., Joubert, M. & Darmoni, S.J. (2012). Aligning biomedical terminologies in French: towards semantic interoperability in medical applications. In *Medical Informatics* (pp. 41–68). New York, NY, USA: InTech.

Merrill, G.H., Ryan, P.B. & Painter, J.L. (2008). Construction and annotation of a UMLS/SNOMED-based drug ontology for observational pharmacovigilance. In *Proceedings of IDAMAP (Intelligent Data Analysis for Biomedicine and Pharmacology)*, Washington, DC, USA.

MeSH (1998). *Medical Subject Headings*. Bethesda, MD, USA: Library of Medicine. Available at: http://www.nlm.nih.gov/mesh/meshhome.html.

Milian, K., Aleksovski, Z., Vdovjak, R., ten Teije, A. & van Harmelen, F. (2010). Identifying disease-centric subdomains in very large medical ontologies: a case-study on breast cancer concepts in SNOMED CT. Or: finding 2500 Out of 300.000. In *Proceedings of the 2009 AIME International Conference on Knowledge Representation for Health-Care: Data, Processes and Guidelines*, Verona, Italy (pp. 50–63). Heidelberg: Springer-Verlag.

Miller, G.A. (1995). WordNet: a lexical database for English. *Communications of the ACM*, *38*(11), 39–41.

Miller, G.A. & Hristea, F. (2006). WordNet nouns: classes and instances. *Computer linguistics*, *32*(1), 1–3.

Mizoguchi, R., Kou, H., Zhou, J., Kozaki, K., Imai, K. & Ohe, K. (2009). An advanced clinical ontology. In *Proceedings of International Conference on Biomedical Ontology*, Buffalo, NY, USA (pp. 119–122).

Moravec, P., Kolovrat, M. & Snásel, V. (2004). LSI vs. WordNet ontology in dimension reduction for information retrieval. In *Proceedings of the Dateso 2004 Annual International Workshop on Databases, Texts, Specifications and Objects*, Desna, Czech Republic (pp. 18–26).

Mörchen, F., Dejori, M., Fradkin, D., Etienne, J., Wachmann, B. & Bundschus, M. (2008). Anticipating annotations and emerging trends in biomedical literature. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Las Vegas, NV, USA (pp. 954–962).

Mottaz, A., Yip, Y.L., Ruch, P. & Veuthey, A.-L. (2008). Mapping proteins to disease terminologies: from UniProt to MeSH. *BMC Bioinformatics*, *9* (Suppl. 5), S3.

Nguyen, A., Lawley, M., Hansen, D. & Colquist, S. (2009). SNOMED clinical terminology mapping from free-text histopathology and radiology reports. In *Proceedings of the COAS–IACR Scientific Meeting*, Sydney, Australia. *Asia–Pacific Journal of Clinical Oncology*, *4* (Suppl. 2), A169.

Niles, I. & Pease, A. (2003). Linking lexicons and ontologies: mapping WordNet to the suggested upper merged ontology. In *Proceedings of the 2003 International Conference on Information and Knowledge Engineering*, Las Vegas, NV, USA (pp. 23–26).

Noy, N.F., de Coronado, S., Solbrig, H., Fragoso, G., Hartel, F.W. & Musen, M.A. (2008). Representing the NCI thesaurus in OWL DL: modeling tools help modeling languages. *Applied Ontology*, *3*(3), 173–190.

Névéol, A., Soualmia, L., Douyère, M., Rogozan, A., Thirion, B. & Darmoni, S.J. (2004). Using CISMeF MeSH "Encapsulated" terminology and a categorization algorithm for health resources. *International Journal of Medical Informatics*, *73*(1), 57–64.

O'Hara, T., Mahesh, K. & Niremburg, S. (1998). Lexical acquisition with WordNet and the microkosmos ontology. In *Proceedings of the ACL Workshop on the Use of WordNet in NLP* (pp. 94–101).

Ongenae, F., Backere, F.D., Steurbaut, K., Colpaert, K., Kerckhove, W., Decruyenaere, J. & Turck, F.D. (2010). Towards computerizing intensive care sedation guidelines: design of a rule-based architecture for automated execution of clinical guidelines. *BMC Medical Informatics and Decision Making*, *10*, 3.

Patel, C.O. & Cimino, J.J. (2006). Mining cross-terminology links in the UMLS. In *Proceedings of the AMIA Annual Symposium* (pp. 624–628).

Patrick, J. (2006). Metonymic and holonymic roles and emergent properties in the SNOMED CT ontology. In *Proceedings of the 2nd Australasian Workshop on Advances in Ontologies*, Hobart, Australia (pp. 61–67). Sydney: Australian Computer Society.

Patrick, J., Wang, Y. & Budd, P. (2007). An automated system for conversion of clinical notes into SNOMED clinical terminology. In *Proceedings of the 5th Australasian Symposium on ACSW Frontiers, ACSW'07*, Ballarat, Australia (Vol. 68, pp. 219–226). Sydney: Australian Computer Society

Patwar, S., Kaur, P., Sylvester, A.G. & Balaji, V. (2009). Towards a novel content organisation in agriculture using semantic technologies: a study with topic maps as a tool. *International Journal of Metadata, Semantics and Ontologies*, *4*(1,2), 65–71.

Pérez-Rey, D., García-Remesal, M., Alonso-Calvo, M., Billhardt, H., Martín-Sánchez, F. & Sousa, A. (2006). ONTOFUSION: ontology-based integration of genomic and clinical databases. *Computers in Biology and Medicine*, *36*, 712–730.

Petrakis, E.G.M., Varelas, G., Hliaoutakis, A. & Raftopoulou, P. (1998). Design and evaluation of semantic similarity measures for concepts stemming from the same or different ontologies. In *Proceedings of the 4th Workshop on Multimedia Semantics (WMS'06)* (pp. 44–52).

Pimentel, J. (2011). Description de verbes juridiques au moyen de la sémantique des cadres. In *Terminologie and ontologie: Théories et applications (TOTH'2011)*, Annecy, France.

Pisanelli, D.M., Gangemi, A. & Steve, G. (1998). An ontological analysis of the UMLS metathesaurus. In *Proceedings of AMIA'98* (pp. 810–814).

Pisanelli, D.M., Gangemi, A. & Steve, G. (1999). A medical ontology library that integrates the UMLS metathesaurus. In *Proceedings of the Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making* (pp. 239–248).

Pustejovsky, J., Rumshisky, A. & Castaño, J.M. (2002). Rerendering semantic ontologies: automatic extensions to UMLS through corpus analytics. In *LREC 2002 Workshop on Ontologies and Lexical Knowledge Bases*, Las Palmas, Canary Islands, Spain.

Qamar, R. & Rector, A. (2007). Semantic mapping of clinical model data to biomedical terminologies to facilitate data interoperability. In *Proceedings of the HealthCare Computing Conference*, Harrogate, UK.

Rak, D., Svatek, V., Fidalgo, M. & Alm, O. (2008). Detecting MeSH keywords and topics in the context of website quality assessment. In *Proceedings of the 1st International Workshop on Describing Medical Web Resources (DRMed'2008)*, Goteborg.

Rector, A.L., Qamar, R. & Marley, T. (2009). Binding ontologies and coding systems to electronic health records and messages. *Applied Ontology*, *4*(1), 51–69.

Rickard, K.L., Mejino Jr., J.L.V., Martin, R.F., Agoncillo, A.V. & Rosse, C. (2004). Problems and solutions with integrating terminologies into evolving knowledge bases. In *Proceedings of MedInfo 2004* (pp. 420–424).

Rosse, C. & Mejino, J.L.V. (2003). A reference ontology for biomedical informatics: the foundational model of anatomy. *Journal of Biomedical Informatics*, *36*(6), 478–500 (Special Issue: Unified Medical Language System).

Schulz, S. & Cornet, R. (2009). SNOMED CT's ontological commitment. In *Proceedings of International Conference on Biomedical Ontology*, Buffalo, NY.

Schulz, S. & Stenzhorn, H. (2007). Ten theses on clinical ontologies. In *Proceedings of the 4th International Council on Medical and Care Compunetics Event*, Amsterdam.

Schulz, S., Suntisrivaraporn, B. & Baader, F. (2007). SNOMED CT's problem list: ontologists' and logicians' therapy suggestions. In *Proceedings of MedInfo 2007*. Studies in Health Technology and Informatics (Vol. 129, pp. 802–806). Amsterdam: IOS Press.

Schulze-Kremer, S., Smith, B. & Kumar, A. (2004). Revising the UMLS semantic network. In M. Fieschi (Ed.), *Proceedings of Medinfo 2004* (p. 1700). Amsterdam: IOS Press.

Schweighofer, E. & Liebwald, D. (2007). Advanced lexical ontologies and hybrid knowledge based systems: first steps to a dynamic legal electronic commentary. *Artificial Intelligence and Law*, *15*(2), 103–115.

Seyed, P., Rector, A., Sattler, U., Parsia, B. & Stevens, R. (2012). Representation of part-whole relationships in SNOMED-CT. In *Proceedings of the International Conference on Biomedical Ontologies*, Graz, Austria.

Simon, J., Fielding, J., Santos, M.D. & Smith, B. (2004). Reference ontologies for biomedical ontology integration and natural language processing. In *Proceedings of EuroMISE 2004*, Prague.

Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., Goldberg, L.J., Eilbeck, K., Ireland, A., Mungall, C.J., Leontis, N., Rocca-Serra, P., Ruttenberg, A., Sansone, S.-A., Scheuermann, R.H., Shah, N., Whetzel, P.L. & Lewis, S. (2007). The OBO foundry: coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnology*, *11*, 1251–1255.

Smith, B., Mejino, J.L.V., Schulz, S., Kumar, A. & Rosse, C. (2005). Anatomical information science. In *Proceedings of the Spatial Information Theory, International Conference*. Lecture Notes in Computer Science (Vol. 3693, pp. 149–164). Berlin/Heidelberg: Springer.

Soergel, D., Lauser, B., Liang, A., Fisseha, F., Keizer, J. & Katz, S. (2004). Reengineering thesauri for new applications: the AGROVOC example. *Journal of Digital Information*, *4*(4), 1–23.

Spackman, K. & Campbell, K. (1998). Compositional concept representation using SNOMED: towards further convergence of clinical terminologies. *Journal of the American Medical Informatics Association*, *5*, 740–744.

Spackman, K.A., Campbell, K.E. & Côté, R.A. (1997). SNOMED RT: a reference terminology for health care. In *Proceedings of the AMIA Annual Fall Symposium* (pp. 640–644).

Spackman, K.A. & Reynoso, G. (2004). Examining snomed from the perspective of formal ontological principles: some preliminary analysis and observations. In *Proceedings of the KR 2004 Workshop on Formal Biomedical Knowledge Representation*, Whistler, BC, Canada (pp. 72–80).

Spasic, I., Ananiadou, S., McNaught, J. & Kumar, A. (2005). Text mining and ontologies in biomedicine: making sense of raw text. *Briefings in Bioinformatics*, *6*(3), 239–251.

Stenzhorn, H., Schulz, S., Boeker, M. & Smith, B. (2008). Adapting clinical ontologies in real-world environments. *Journal of Universal Computer Science*, *14*(22), 3767–3780.

Suchanek, F.M., Kasneci, G. & Weikum, G. (2008). YAGO: a large ontology from Wikipedia and WordNet. *Web Semantics: Science, Services and Agents on the World Wide Web*, *6*(3), 203–217.

van Assem, M., Malaisé, V., Miles, A. & Schreiber, G. (2006). A method to convert thesauri to SKOS. In Y. Sure and J. Domingue (Eds.), *The Semantic Web: Research and Applications*. Lecture Notes in Computer Science (Vol. 4011, pp. 95–109). Berlin/Heidelberg: Springer.

van Rees, R. (2003). Clarity in the usage of the terms ontology, taxonomy and classification. *CIB Report*, *284*, 432–439.

Vintar, S., Todorovski, L., Sonntag, D. & Buitelaar, P. (2003). Evaluating context features for medical relation mining. In *Proceedings of the ECML/PKDD Workshop on Data Mining and Text Mining for Bioinformatics*, Cavtat-Dubrovnik, Croatia.

Wong, S.H.S. (2004). Fighting arbitrariness in wordnet-like lexical databases – a natural language motivated remedy. In *Proceedings of GWC 2004, the 2nd Global WordNet Conference*, Masaryk University, Brno (pp. 234–241).

Wu, S.T., Liu, H., Li, D., Tao, C., Musen, M.A., Chute, C.G. & Shah, N.H. (2012). Unified medical language system term occurrences in clinical notes: a large-scale corpus analysis. *Journal of the American Medical Informatics Association*, *19*, e149–e156.

Yoo, I. & Hu, X. (2006). Biomedical ontology MeSH improves document clustering qualify on MEDLINE articles: a comparison study. In *Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems* (pp. 577–582). Washington, DC, USA: IEEE Computer Society.

Yu, A.C. (2006). Methods in biomedical ontology. *Journal of Biomedical Informatics*, *39*(3), 252–266.

Zhang, S. & Bodenreider, O. (2003). Knowledge augmentation for aligning ontologies: an evaluation in the biomedical domain. In *Proceedings of the Semantic Integration Workshop at the Second International Semantic Web Conference (ISWC'2003)* (pp. 109–114). Available at: http://mor.nlm.nih.gov/pubs/pdf/2003-iswc-semint-sz.pdf.