

Over the past decades, ontologies have become key components of information systems [1, 2] and have found various applications such as natural language processing [3], software engineering [4] and knowledge management in the Semantic Web [5]. The use of ontologies also spans many domains including biomedicine, geography and e-commerce. In the biomedical domain – our focus in this special issue – ontologies play an important role in research [6], healthcare [7] and translational biomedicine [8], supporting tasks including knowledge management, data integration and decision support [9].

A large number of biomedical ontologies have been created by individual researchers, academic consortia, institutions and companies. Most ontologies are created for a given purpose (e.g., to annotate gene products) and independently of existing ontologies. As a consequence, although ontologies are ideally created for sharing and reusing knowledge [10], their impact can be limited to a small audience by lack of diffusion.

The abundance of ontologies should be good news to biomedical researchers, as they can in theory select the best ontology for their application. In practice, however, there is no comprehensive registry of biomedical ontologies and the few existing collections of ontologies do not provide enough information about the ontologies to support effective discoverability.

In the biomedical domain, the oldest repository is the Unified Medical Language System (UMLS) [11, 12], developed at the National Library of Medicine, which currently integrates 148 biomedical terminologies and ontologies. More recently, the National Center for Biomedical Ontologies (NCBO) created the BioPortal, a web-based application providing access to 129 ontologies in the biomedical domain [13, 14]. Analogously, the Ontology Lookup Service [15, 16] developed by the European Bioinformatics Institute supports queries against 61 ontologies from the Open Biomedical Ontology (OBO) family¹. These three resources offer online lookup services, i.e., support the discovery of individual entities through their names and identifiers. While generally useful to researchers, none of these resources is comprehensive, and interfaces to ontologies are no substitute for a rich description of their content and characteristics.

The proliferation of knowledge bases has led to the creation of registries such as Oxford University Press' Nucleic Acids Research (NAR) "database issue" [17]. With both a print edition (available under the open access model) and a companion online version (the Molecular Biology Database Collection), NAR provides a description of a large number of electronic biomedical resources. The databases listed in the print edition (182 in the 2008) include *GenBank*, *UniProt*, *KEGG*, the *Stanford Tissue Microarray Database* and *DrugBank*. In addition, the online edition offers a more succinct description of over one thousand resources. The descriptions contributed by the authors of these resources are curated by NAR through their usual editorial process. By providing a textual description indexed in bibliographic databases such as PubMed/Medline (print version) and a set of metadata (online version), the NAR database edition acts as a registry of biomedical resources and contributes to their discoverability.

No such registry currently exists for ontologies, neither for biomedicine, nor for other domains. The 2008 edition of the Ontology Summit only laid the blueprints of an Open Ontology Repository [18], which will provide a framework for implementing such registries. The definition of metadata for ontologies is an

¹ The OBO Foundry also serves as a repository of ontologies from the OBO family.

active field of research and initiatives such as the Ontology Metadata Vocabulary (OMV) are attempts at providing for ontologies the standard metadata offered by the Dublin Core for text documents [19].

This Applied Ontology special issue reports on the first results of the MetaOntology initiative, developed in partnership with the National Center for Biomedical Ontology (NCBO), one of the seven National Centers for Biomedical Computing in the U.S. In addition to being a repository for biomedical ontologies, NCBO's BioPortal [14] "defines relationships among those ontologies and between the ontologies and online data resources" and supports "community-based participation in the evaluation and evolution of ontology content" [13]. In other words, BioPortal enables biomedical ontology users to discover, visualize, download and evaluate existing resources and their interrelations. BioPortal will host the ontologies from the MetaOntology issue and their associated metadata, as well as links to the articles in Applied Ontology.

Current metadata about these ontologies and tools include the name, type and version of each resource, as well as the URI where it can be downloaded and contact information. Additionally for ontologies, we collect the number of entities, relations and instances represented, as well as information about the domain and formalism used for the representation. For tools, the purpose and computer platform are recorded. The complete list of metadata requested from ontology developers can be found in the call for papers of the MetaOntology [20].

Participating in the MetaOntology gives the authors of biomedical ontology resources (ontologies and tools) an opportunity to publicize their resources to both ontology specialists (through Applied Ontology) and domain experts (through BioPortal). The print version provides a bibliographic reference that can be cited by other researchers when they use a given resource as part of their projects. The online version contributes to the discoverability, diffusion and reuse of these resources. Ontologies integrated in the BioPortal become immediately visible and can be browsed and downloaded. Additionally, authors receive feedback from the community through the collaborative evaluation fostered by BioPortal.

This inaugural edition of the MetaOntology for the biomedical domain lists five resources. Three of these – *BioTop*, *BioZen* and *GFO-Bio* – are top domain ontologies, i.e., ontologies in reference to which domain ontologies can be created. The other two are *OntoPneumo*, an ontology of pneumology developed to support coding, and one integrative resource combining several pathway ontologies. With five ontologies described, the current edition of the MetaOntology is still at an embryonic stage. As it provides a tribunal for authors to publicize their ontologies, as well as a resource for potential users to discover them, we expect the MetaOntology to grow rapidly and eventually expand beyond the confines of biomedicine. To achieve such goal, further contributions are expected concerning the proper ways to organize, compare, catalog and retrieve domain ontologies.

References

1. Gruber, T.R.: A Translation Approach to Portable Ontology Specifications. *Knowl. Acquis.* 5, 199-220 (1993)
2. Guarino, N.: Formal Ontology in Information Systems. In: Guarino, N. (ed.): *Formal Ontology in Information Systems: Proceedings of the 1st International Conference June 6-8, 1998, Trento, Italy*, Vol. 46, pp. 3-15. IOS Press, Amsterdam (1998)
3. Nirenburg, S., Raskin, V.: *Ontological semantics*. MIT Press, Cambridge, Mass. (2004)
4. Calero, C., Ruiz, F., Piattini, M.: *Ontologies for software engineering and software technology*. Springer, Berlin ; New York (2006)
5. Kashyap, V.: *The semantic web : semantics for data and services on the web*. Springer, New York (2008)
6. Blake, J.A., Bult, C.J.: Beyond the data deluge: data integration and bio-ontologies. *J Biomed Inform* 39, 314-320 (2006)
7. Garde, S., Knaup, P., Hovenga, E., Heard, S.: Towards semantic interoperability for electronic health records. *Methods Inf Med* 46, 332-343 (2007)
8. Ruttenberg, A., Clark, T., Bug, W., Samwald, M., Bodenreider, O., Chen, H., Doherty, D., Forsberg, K., Gao, Y., Kashyap, V., Kinoshita, J., Luciano, J., Marshall, M.S., Ogbuji, C., Rees, J., Stephens, S., Wong, G.T., Wu, E., Zaccagnini, D., Hongsermeier, T., Neumann, E., Herman, I., Cheung, K.H.: Advancing translational research with the Semantic Web. *BMC Bioinformatics* 8 Suppl 3, S2 (2007)
9. Bodenreider, O.: Biomedical ontologies in action: role in knowledge management, data integration and decision support. Geissbuhler A, Kulikowski C, editors. *IMIA Yearbook of Medical Informatics 2008*. *Methods Inf Med* 47, 67-79 (2008)
10. Musen, M.A.: Dimensions of knowledge sharing and reuse. *Comput Biomed Res* 25, 435-467 (1992)
11. Bodenreider, O.: The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Res* 32, D267-270 (2004)
12. UMLS Knowledge Source Server, <http://umlsks.nlm.nih.gov/>
13. Musen, M., Shah, N., Noy, N., Dai, B., Dorf, M., Griffith, N., Buntrock, J.D., Jonquet, C., Montegut, M.J., Rubin, D.L.: BioPortal: Ontologies and Data Resources with the Click of a Mouse. *AMIA Annu Symp Proc*, 1223-1224 (2008)
14. BioPortal, <http://bioportal.bioontology.org/>
15. Cote, R.G., Jones, P., Martens, L., Apweiler, R., Hermjakob, H.: The Ontology Lookup Service: more data and better tools for controlled vocabulary queries. *Nucleic Acids Res* 36, W372-376 (2008)
16. Ontology Lookup Service, <http://www.ebi.ac.uk/ontology-lookup/>
17. Bateman, A.: Editorial. *Nucl. Acids Res.* 36, D1- (2008)
18. Towards an Open Ontology Repository, http://ontolog.cim3.net/cgi-bin/wiki.pl?OntologySummit2008_Communique
19. Ontology Metadata Vocabulary, <http://omv.ontoware.org/>
20. Call for Submissions to the Applied Ontology MetaOntology Issue: Biomedical Ontologies. *Applied Ontology* 1, 213-215 (2006)