JOURNAL OF
BIOMEDICAL SEMANTICS

INTRODUCTION   Open Access

OBML – Ontologies in Biomedicine and Life Sciences

Heinrich Herre[1*], Robert Hoehndorf[2], Janet Kelso[3], Frank Loebe[4], Stefan Schulz[5]

* Correspondence: heinrich.
herre@imise.uni-leipzig.de
[1]Institute for Medical Informatics,
Statistics and Epidemiology (IMISE),
University of Leipzig, Haertelstrasse
16-18, 04107 Leipzig, Germany

Abstract

The OBML 2010 workshop, held at the University of Mannheim on September 9-10, 2010, is the 2nd in a series of meetings organized by the Working Group "Ontologies in Biomedicine and Life Sciences" of the German Society of Computer Science (GI) and the German Society of Medical Informatics, Biometry and Epidemiology (GMDS). Integrating, processing and applying the rapidly expanding information generated in the life sciences — from public health to clinical care and molecular biology — is one of the most challenging problems that research in these fields is facing today. As the amounts of experimental data, clinical information and scientific knowledge increase, there is a growing need to promote interoperability of these resources, support formal analyses, and to pre-process knowledge for further use in problem solving and hypothesis formulation.

The OBML workshop series pursues the aim of gathering scientists who research topics related to life science ontologies, to exchange ideas, discuss new results and establish relationships. The OBML group promotes the collaboration between ontologists, computer scientists, bio-informaticians and applied logicians, as well as the cooperation with physicians, biologists, biochemists and biometricians, and supports the establishment of this new discipline in research and teaching. Research topics of OBML 2010 included medical informatics, Semantic Web applications, formal ontology, bio-ontologies, knowledge representation as well as the wide range of applications of biomedical ontologies to science and medicine. A total of 14 papers were presented, and from these we selected four manuscripts for inclusion in this special issue.

An interdisciplinary audience from all areas related to biomedical ontologies attended OBML 2010. In the future, OBML will continue as an annual meeting that aims to bridge the gap between theory and application of ontologies in the life sciences. The next event emphasizes the special topic of the *ontology of phenotypes*, in Berlin, Germany on October 6-7, 2011.

Introduction

Utilizing the rapidly growing body of data and knowledge in the life sciences poses a severe challenge today. Integration of data and knowledge is particularly important in translating research outputs from molecular biology and genetics into improvements in the treatments of disease. A prime goal of genetics, molecular biology and biochemistry is to reveal the relation between molecular mechanisms and their resulting contribution to an organism's phenotype. Recent advances in high-throughput sequencing technologies have facilitated the rapid identification of the genomes of a range of individual organisms. The development of high-throughput technologies in molecular biology has enabled the possibility for a paradigm shift towards the personalized treatment of disease based on an individual patient's genetic markup. To facilitate this paradigm shift, biological knowledge about the disease mechanisms must be accessible to scientific and clinical analyses. In particular the molecular, genetic and phenotypic information that is relevant to understanding human disease can be investigated in organisms of different species that may serve as models for humans.

The challenge is to represent, compare and analyze information within and across domains and representation formats, and resulting from various analytical methods in order to produce knowledge bases that are amenable to scientific investigation and clinical application. Ontologies provide formal specifications and harmonized definitions of representational units (types, classes, concepts) and thus provide the building blocks for the computable modeling of domain knowledge. Methods in formal ontology support and enhance the semantic foundation, the standardization and the analysis of concepts related to a domain of reality and underpin the correct formalization of domain knowledge. Additionally, ontologies play an important role in the context of the Semantic Web promoting new paradigms of knowledge representation, processing and distribution.

Integration of data and knowledge in science benefits from a precise and formal characterization of data that is based on ontologies and the use of symbolic logic. In contrast to heuristic, approximate or statistical approaches, methods based on expressive logics provide a justification and explanation of the answers they provide and deliver a degree of certainty that is not achievable through other methods. They are therefore particularly suitable for the representation of scientific knowledge and can provide a layer based on which biomedical information can be verified, exchanged and integrated.

Ontologies also provide a rich graph-structure, which can be dynamically generated through automated inference, and which can be utilized to improve statistical methods or enable the use of distance metrics to measure semantic similarity. Ontologies based on expressive logics can further provide the means to verify results of scientific analyses, reject contradictory hypotheses and reveal the implications of scientific findings.

The OBML workshop series pursues the aim of gathering scientists who investigate topics related to life science ontologies, who exchange ideas, discuss new results and promote cooperations and collaborations. Research topics of OBML 2010 included medical informatics, Semantic Web applications, formal ontology, bio-ontologies, knowledge representation as well as the wide range of applications of biomedical ontologies to science and medicine. Out of this wide range of topics, we selected four manuscripts, which are contained in this special issue.

Contributions to OBML 2010

Extended articles in this issue

In the area of knowledge representation and applied ontology, the OBML workshop benefited from discussions of how to represent dispositions [1] and grains, components and mixtures [2] in biomedical ontologies. These topics are important for representing diseases, chemical entities and physiological processes that involve solutions such as blood. Dispositions are widely discussed both in philosophy [3-6], in the formal ontology community [7-10] as well as in the context of medical applications and representing disease [11,12]. A major challenge in representing dispositions is their context dependence, their dependence on a trigger, as well as the relation between dispositions and the processes that may be their realizations. Another problem arises due to the absence of relations with more than two arguments in description logics underlying the Web Ontology Language (OWL) [13,14]. The article "Representing dispositions" investigates how to represent dispositions in OWL, including their triggers and contexts, discusses applications in disease ontologies and ontologies of chemistry and provides an outlook on complex dispositions that can be realized through multiple kinds of processes or use multiple triggers. The paper "Grains, Components and Mixtures in Biomedical Ontologies" discusses competing representations of collectives (multiple entities of the same sort) for an adequate representation of substance mixtures. The authors argue that a strict dichotomy between (homogeneous) collectives and (heterogeneous) compounds, as advocated in previous work, is problematic. For instance, the distinction between isomeric subtypes of a molecule can be important in one use case but might be neglected in another one. Two different ways of representing mixtures are presented: (i) mixtures are the additions of two or more fractions of collectives, or (ii), mixtures are collectives of two or more types of granular parts. Using OWL-DL and the upper-level ontology BioTop [15], the authors demonstrate how the equivalence between both representations can be computed.

The article "Anatomy Ontologies and Potential Users: Bridging the Gap" evaluates the overlap of terms in anatomy ontologies and provides insight in how to align species-specific anatomies into a cross-species anatomy ontology. For this purpose, the authors evaluated how well the class names in anatomy ontologies fit the terms that are used by annotators. The authors selected three comprehensive data sets of annotations and evaluated the Foundational Model of Anatomy (FMA) [16] and the species-independent UBERON anatomy ontology [17] using lexical matching approaches. As a result, the authors identified a mismatch between the terms used by annotators and the terms provided by anatomy ontologies. These constitute short-comings in anatomy ontologies, which should be addressed by their developers to bridge the gap between the ontologies' formal representation and their potential users.

Clinical applications of ontologies are the topic of the article titled "An Ontologically Founded Basic Architecture for Information Systems in Clinical and Epidemiological Research". The authors present an architecture for clinical information systems founded in the General Formal Ontology (GFO) [18] and demonstrate applications to the representation of phenotypes as well as to clinical trials. Since the architecture is based on a top-level ontology, it is argued that such a meta-model can be applied in several scenarios, including clinical trial data and epidemiological research.

Further contributions at the workshop

Several presentations covering a broad range of topics were given at the OBML 2010 workshop. The use of a semantic wiki, SBML2SMW, which is based on the Semantic MediaWiki architecture [19] and applied to the Systems Biology Markup Language [20] for application to the domain of systems biology was demonstrated [21]. SBML2SMW establishes a bi-directional information flow between a semantic wiki system and the domain of systems biology, and thereby enables the use of automated reasoning to discover hidden knowledge in biosimulation models.

Another demonstration of implementing ontology-based software in health care was given by C. Cocos and W. MacCaull [22], who illustrated how the Basic Formal Ontology [23] and its theory of roles can contribute to the implementation of access control policies in a healthcare information system.

Zaveri et al. [24] used the Linked OpenData cloud to evaluate the disparity between active areas of research and the global burden of disease. The presentation demonstrated a research program to automatically identify research areas and links between them, and illustrated potential visualization as well as analysis methods.

An important topic at OBML 2010 was knowledge representation, reasoning and foundations of biomedical ontologies, and several manuscripts contributed insights into these areas of knowledge. N. Grewe [25] provided a comprehensive discussion of strategies on representing n-ary relations in OWL. The topic of n-ary relations has applications from representing dispositions over the representation of processes to relations that contain temporal arguments.

The question whether anatomic cavities can be inflamed (e.g., by a verbatim interpretation of the term "Sinusitis"), or whether it must be the material boundary of a cavity such as a nasal sinus, was raised by J. Niggemann et al. [26], who illustrated the relevance of such questions for SNOMED CT concept definitions and the reliability of inferences.

Further manuscripts presented at the OBML workshop discussed the construction of an ontology for primary immuno-deficiencies [27], a cellular genealogical tree ontology [28], strategies for improving phenotype ontologies [29], the importance of simplifying ontologies to improve comprehensibility and understanding [30], as well as practical decisions on pre- vs. post-coordination in ontology engineering [31].

Outlook

In 2010, the OBML workshop could gather an interdisciplinary audience from all areas related to biomedical ontologies. In the future, OBML will continue as a yearly event, organized by the working group, that aims to bridge the gap between theory and application of ontologies in the life sciences. To achieve this goal, the OBML working group will collaborate with the International Association for Ontology and its Applications (IAOA), the German Society of Computer Science (GI), and the German Society of Medical Informatics, Biometry and Epidemiology (GMDS). The OBML organizing and program committees will span all scientific disciplines associated with ontologies: from knowledge representation, logics and philosophy over Semantic Web technologies, natural language processing and statistics to applications in all systems, domains and levels of granularity studied in biomedicine.

The next OBML workshop will take place in Berlin on October 6-7, 2011, and it will be partly dedicated to the special topic of the *ontology of phenotypes.*

Author details

[1]Institute for Medical Informatics, Statistics and Epidemiology (IMISE), University of Leipzig, Haertelstrasse 16-18, 04107 Leipzig, Germany. [2]Department of Genetics, Downing Street, CB2 3EH Cambridge, UK. [3]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany. [4]Department of Computer Science, University of Leipzig, Johannisgasse 26, 04103 Leipzig, Germany. [5]Institute of Medical Informatics, Statistics and Documentation, Medical University of Graz, Auenbruggerplatz 2, 8036 Graz, Austria.

Competing interests

The authors declare that they have no competing interests.

Published: 9 August 2011

References

1. Röhl J, Jansen L: **Representing Dispositions.** In *OBML 2010 Workshop Proceedings* Herre H, Hoehndorf R, Kelso J, Schulz S, Institut fuer Medizinische Informatik, Statistik und Epidemiologie (IMISE), Markus Loeffler 2010, F1-5[http://www.onto-med.de/obml/ws2010/obml2010report.pdf].
2. Jansen L, Schulz S: **Grains, Components and Mixtures in Biomedical Ontologies.** In *OBML 2010 Workshop Proceedings* Herre H, Hoehndorf R, Kelso J, Schulz S, Institut fuer Medizinische Informatik, Statistik und Epidemiologie (IMISE), Markus Loeffler 2010, G1-4[http://www.onto-med.de/obml/ws2010/obml2010report.pdf].
3. Armstrong DM: **A World of States of Affairs (Cambridge Studies in Philosophy).** Cambridge University Press; 1997.
4. Choi S: **The Simple Vs. Reformed Conditional Analysis of Dispositions.** *Synthese* 2006, **148(2)**:369-379[http://dx.doi.org/10.1007/s11229-004-6229-z].
5. Lewis D: **Finkish Dispositions.** *The Philosophical Quarterly* 1997, **47(187)**:143-158[http://dx.doi.org/10.1111/1467-9213.00052].
6. Johansson I: **Functions, function concepts, and scales.** *The Monist* 2004, **87**:96-114.
7. Hoehndorf R, Kelso J, Herre H: **Contributions to the formal ontology of functions and dispositions: An application of non-monotonic reasoning.** *Proceedings of Bio-Ontologies 2009: Knowledge in Biology* 2009.
8. Arp R, Smith B: **Function, Role, and Disposition in Basic Formal Ontology.** *Proceedings of The 11th Annual Bio-Ontologies Meeting* 2008.
9. Batchelor C, Hastings J, Steinbeck C: **Ontological dependence, dispositions and institutional reality in chemistry.** In *Formal Ontology in Information Systems - Proceedings of the Sixth International Conference (FOIS 2010). Volume 209.* IOS Press;Galton A, Mizoguchi R 2010:271-284.
10. Hoehndorf R, Oellrich A, Dumontier M, Kelso J, Rebholz-Schuhmann D, Herre H: **Relations as patterns: Bridging the gap between OBO and OWL.** *BMC Bioinformatics* 2010, **11**:441+.
11. Goldfain A, Smith B, Cowell L: **Dispositions and the Infectious Disease Ontology.** *Proceedings of Formal Ontologies in Information Systems (FOIS)* 2010, 400-413.
12. Schulz S, Stenzhorn H, Boeker M, Smith B: **Strengths and limitations of formal ontologies in the biomedical domain.** *RECIIS Rev Electron Comun Inf Inov Saude* 2009, **3(1)**:31-45.
13. Grau B, Horrocks I, Motik B, Parsia B, Patelschneider P, Sattler U: **OWL 2: The next step for OWL.** *Web Semantics: Science, Services and Agents on the World Wide Web* 2008, **6(4)**:309-322[http://dx.doi.org/10.1016/j.websem.2008.05.001].
14. Horrocks I, Kutz O, Sattler U: **The Even More Irresistible SROIQ.** In *KR.* AAAI Press;Doherty P, Mylopoulos J, Welty CA 2006:57-67[http://dblp.uni-trier.de/db/conf/kr/kr2006.html#HorrocksKS06].
15. Schulz S, Beisswanger E, Wermter J, Hahn U: **Towards an upper level ontology for molecular biology.** *AMIA Annu Symp Proc* 2006, 694-698[http://view.ncbi.nlm.nih.gov/pubmed/17238430].
16. Rosse C, Mejino JLV: **A Reference Ontology for Biomedical Informatics: the Foundational Model of Anatomy.** *Journal of Biomedical Informatics* 2003, **36(6)**:478-500[http://dx.doi.org/10.1016/j.jbi.2003.11.007].
17. Mungall C, Gkoutos G, Smith C, Haendel M, Lewis S, Ashburner M: **Integrating phenotype ontologies across multiple species.** *Genome Biology* 2010, **11**:R2+[http://dx.doi.org/10.1186/gb-2010-11-1-r2].
18. Herre H: **General Formal Ontology (GFO): A Foundational Ontology for Conceptual Modelling.** In *Theory and Applications of Ontology: Computer Applications.* Heidelberg: Springer;Poli R, Healy M, Kameas A 2010:297-345[http://dx.doi.org/10.1007/978-90-481-8847-5_14].
19. Krötzsch M, Vrandečić D, Völkel M, Haller H, Studer R: **Semantic Wikipedia.** *Web Semantics: Science, Services and Agents on the World Wide Web* 2007, **5(4)**:251-261.
20. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, Cuellar AA, Dronov S, Gilles ED, Ginkel M, Gor V, Goryanin II, Hedley WJ, Hodgman TC, Hofmeyr JH, Hunter PJ, *et al*: **The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models.** *Bioinformatics* 2003, **19(4)**:524-531[http://bioinformatics.oxfordjournals.org/content/19/4/524.abstract].
21. Mathäss T, Haase P, Kitano H, Toldo L: **SBML2SMW: Bridging System Biology with Semantic Web Technologies for Biomedical Knowledge Acquisition and Hypothesis Elicitation.** In *OBML 2010 Workshop Proceedings* Herre H,

Hoehndorf R, Kelso J, Schulz S, Institut fuer Medizinische Informatik, Statistik und Epidemiologie (IMISE), Markus Loeffler 2010, A1-4[http://www.onto-med.de/obml/ws2010/obml2010report.pdf].

22. Cocos C, MacCaull W: **An Ontological Implementation of a Role-Based Access Control Policy for Health Care Information.** In *OBML 2010 Workshop Proceedings* Herre H, Hoehndorf R, Kelso J, Schulz S, Institut fuer Medizinische Informatik, Statistik und Epidemiologie (IMISE), Markus Loeffler 2010, B1-5[http://www.onto-med.de/obml/ws2010/obml2010report.pdf].

23. Grenon P: **BFO in a Nutshell: A Bi-categorial Axiomatization of BFO and Comparison with DOLCE.** *IFOMIS Report 06/2003* Institute for Formal Ontology and Medical Information Science (IFOMIS), University of Leipzig, Leipzig, Germany; 2003 [http://www.ifomis.uni-saarland.de/Research/IFOMISReports/IFOMIS%20Report%2006_2003.pdf].

24. Zaveri A, Pietrobon R, Ermilov T, Martin M, Heino N, Auer S: **Evaluating the Disparity between Active Areas of Biomedical Research and the Global Burden of Disease Employing Linked Data and Data-driven Discovery.** In *OBML 2010 Workshop Proceedings* Herre H, Hoehndorf R, Kelso J, Schulz S, Institut fuer Medizinische Informatik, Statistik und Epidemiologie (IMISE), Markus Loeffler 2010, C1-7[http://www.onto-med.de/obml/ws2010/obml2010report.pdf].

25. Grewe N: **A Generic Reification Strategy for n-ary Relations in DL.** In *OBML 2010 Workshop Proceedings* Herre H, Hoehndorf R, Kelso J, Schulz S, Institut fuer Medizinische Informatik, Statistik und Epidemiologie (IMISE), Markus Loeffler 2010, N1-5[http://www.onto-med.de/obml/ws2010/obml2010report.pdf].

26. Niggemann J, Straub HR, Herre H: **Can a Hole Be Inflamed? On the Relation of Morphologic Abnormalities and Anatomical Cavities in SNOMED CT.** In *OBML 2010 Workshop Proceedings* Herre H, Hoehndorf R, Kelso J, Schulz S, Institut fuer Medizinische Informatik, Statistik und Epidemiologie (IMISE), Markus Loeffler 2010, E1-4[http://www.onto-med.de/obml/ws2010/obml2010report.pdf].

27. Adams N, Hennig C, Hoehndorf R, Oellrich A, Rebholz-Schuhmann D, Hansen G: **The Ontology of Primary Immunodeficiency Diseases (PIDs) - Using PIDs to Rethink the Ontology of Phenotypes.** In *OBML 2010 Workshop Proceedings* Herre H, Hoehndorf R, Kelso J, Schulz S, Institut fuer Medizinische Informatik, Statistik und Epidemiologie (IMISE), Markus Loeffler 2010, I1-4[http://www.onto-med.de/obml/ws2010/obml2010report.pdf].

28. Burek P, Herre H, Röder I, Glauche I, Scherf N, Löffler M: **Towards a Cellular Genealogical Tree Ontology.** In *OBML 2010 Workshop Proceedings* Herre H, Hoehndorf R, Kelso J, Schulz S, Institut fuer Medizinische Informatik, Statistik und Epidemiologie (IMISE), Markus Loeffler 2010, K1-5[http://www.onto-med.de/obml/ws2010/obml2010report.pdf].

29. Oellrich A, Rebholz-Schuhmann D: **A Classification of Existing Phenotypical Representations and Methods for Improvement.** In *OBML 2010 Workshop Proceedings* Herre H, Hoehndorf R, Kelso J, Schulz S, Institut fuer Medizinische Informatik, Statistik und Epidemiologie (IMISE), Markus Loeffler 2010, J1-4[http://www.onto-med.de/obml/ws2010/obml2010report.pdf].

30. Schober D, Boeker M: **Ontology Simplification: New Buzzword or Real Need?** In *OBML 2010 Workshop Proceedings* Herre H, Hoehndorf R, Kelso J, Schulz S, Institut fuer Medizinische Informatik, Statistik und Epidemiologie (IMISE), Markus Loeffler 2010, M1-5[http://www.onto-med.de/obml/ws2010/obml2010report.pdf].

31. Schulz S, Schober D, Raufie D, Boeker M: **Pre- and Postcoordination in Biomedical Ontologies.** In *OBML 2010 Workshop Proceedings* Herre H, Hoehndorf R, Kelso J, Schulz S, Institut fuer Medizinische Informatik, Statistik und Epidemiologie (IMISE), Markus Loeffler 2010, L1-4[http://www.onto-med.de/obml/ws2010/obml2010report.pdf].