Proceedings of 2004 the Pacific Symposium on Biocomputing. World Scientific; 2003. p. 250-261.
© 2003 World Scientific Publishing Co. Pte. Ltd.

BIOMEDICAL ONTOLOGIES

O. BODENREIDER

U.S. National Library of Medicine
8600 Rockville Pike, MS 43, Bethesda, Maryland, 20894, USA
E-mail: olivier@nlm.nih.gov

J. A. MITCHELL

University of Missouri, Department of Health Management & Informatics,
Columbia, Missouri, 65211, USA
E-mail: MitchellJo@health.missouri.edu

A. T. MCCRAY

U.S. National Library of Medicine
8600 Rockville Pike, MS 52, Bethesda, Maryland, 20894, USA
E-mail: mccray@nlm.nih.gov

As we celebrate the 50th anniversary of the description of the structure of DNA, biology is evolving from a science of organisms and molecules to a science of information. In modern biology, massive amounts of data are produced resulting, for example, from sequencing the genome of many organisms and studying gene expression under various conditions. In turn, there has been a shift from hypothesis-driven experiments to data-driven experiments. Ontologies provide a conceptualization of a domain that can be shared among diverse groups of researchers and health care professionals and used computationally for multiple purposes. Biologic knowledge is evolving so rapidly that it is difficult for most scientists to assimilate and integrate the new information with their existing knowledge. Promoting the creation and use of ontologies for the field and linking to other ontologies in related domains holds the promise of assisting those working in biomedical disciplines and thus making more rapid scientific progress.

The papers presented in this session reflect the ontological needs arising in the biomedical community: sharing the experience of ontology developers and users on the one hand, and developing methods for auditing and evaluating existing ontologies and formalisms, as well as for assessing the usefulness of ontologies in biological applications on the other.

Three papers focus on building, using, and aligning ontologies in various subdomains of biomedicine. Mouse phenotype ontologies are the object of two of these papers. One reports on building ontologies for mouse phenotypes based on the

Phenotype and Trait Ontology (PATO) schema. The other presents the mapping of Phenoslim, another mouse phenotype ontology, to clinical terminologies (UMLS® and SNOMED-CT®). The last paper in this series reports on creating a hierarchy of evidence codes and discusses its application to pathway databases.

One trend in this session is the analysis of the limitations of the formalisms currently used for representing ontologies, with two papers focusing on two different formalisms. The first one analyses issues in the representation of anatomy in ontologies built on the model of Gene Ontology™ (GO). The representation of defaults and exceptions in the Ontology Web Language (OWL) framework is investigated in the second one.

The next two papers focus on making explicit the ontological relations embedded in concept names, thus providing additional auditing methods for these ontologies. One paper investigates implicit knowledge in the Foundational Model of Anatomy and GALEN and its applications to auditing and aligning ontologies. The other paper analyzes concept names nested within GO terms from which ontological relations can be acquired and discusses the application of this method to auditing GO structure.

Finally, the last paper proposes methods for analyzing ambiguity in gene names and its consequences in information extraction.

Although the papers selected for this PSB 2004 session on Biomedical Ontologies may not be representative of all ongoing research efforts in the community, we believe that these papers characterize important research directions in this field. Ontologies need to move from loosely organized sets of terms to frameworks supported by formal properties. The limitations of the formalisms used to represent ontolgies need to be carefully identified and studied. Finally, the current focus on ontologies of anatomy is not surprising since anatomy – from macroscopic to subcellular structures – is a core subdomain of biomedicine whose representation is needed in virtually any biomedical application.