Chapter 8.2
A Measurement Ontology Generalizable for Emerging Domain Applications on the Semantic Web

Henry M. Kim
York University, Canada

Arijit Sengupta
Wright State University, USA

Mark S. Fox
University of Toronto, Canada

Mehmet Dalkilic
Indiana University, USA

ABSTRACT

This article introduces a measurement ontology for applications to Semantic Web applications, specifically for emerging domains such as microarray analysis. The Semantic Web is the next-generation Web of structured data that are automatically shared by software agents, which apply definitions and constraints organized in ontologies to correctly process data from disparate sources. One facet needed to develop Semantic Web ontologies of emerging domains is creating ontologies of concepts that are common to these domains. These general "common-sense" ontologies can be used as building blocks to develop more domain-specific ontologies. However most measurement ontologies concentrate on representing units of measurement and quantities, and not on other measurement concepts such as sampling, mean values, and evaluations of quality based on measurements. In this article, we elaborate on a measurement ontology that represents all these concepts. We present the generality of the ontology, and describe how it is developed, used for analysis and validated.

INTRODUCTION

According to Tim Berners-Lee, whom many attribute as the inventor of the World Wide Web, the Web will evolve into the *Semantic Web*, which relies upon using machine processable domain knowledge represented in *ontologies* to execute and compose automated *Web services* (Berners-Lee, Hendler, & Lassila, 2001; Chen, Zhou, & Zhang, 2006). An ontology is a data model that "consists of a representational vocabulary with precise definitions of the meanings of the terms of this vocabulary plus a set of formal axioms that constrain interpretation and well-formed use of these terms" (Campbell & Shapiro, 1995). Ontology use ensures that data instances are so precisely defined and constrained that the instances can be processed automatically and accurately by Web-based computer programs, or *software agents*. Berners-Lee's et. al (2001) vision of the Semantic Web is that "many software agents, accessing data instances and applying ontologies to the instances, execute Web services in concert, where agents, data instances, and ontologies are distributed all over the Web."

Ontologies for the Semantic Web represent an emerging method for modeling the semantics required to interpret data. In a similar vein, applications such as genomics and GIS's represent emerging domains represented for semantic modeling. There exist Semantic Web ontologies for traditional applications in computer science, and business (Davies, Duke, & Stonkus, 2002; Gandon & Sadeh, 2004; Klischewski & Jeenicke, 2004). There are also emerging domains modeled using traditional semantics modeling techniques (Khatri, Ram, & Snodgrass, 2004; Ram & Wei, 2004). There are even some ontologies of emerging domains such as representations of the Gene Ontology (Ashburner et al., 2000; Wroe, Stevens, Goble, & Ashburner, 2003) and a bioinformatics ontology (Stevens, Goble, Horrocks, & Bechhofer, 2002), which are represented in the de facto Semantic Web Ontology Language, OWL (McGuinness & van Harmelen, 2003), or its predecessor, DAML+OIL (Bechhofer, Goble, & Horrocks, 2001). Although one of the goals of ontology development is a generalization of terms in an application, it is possible to make intelligent choices when several ontologies are available for the same domain (Lozano-Tello & Gomez-Perez, 2004).

When contemplating the development of ontologies of any domain, it is instructive to state the following informal definition: "an ontology is an explicit representation of shared understanding" (Gruber, 1993). Gruber also outlines the conundrum of ontological commitment: the more one commits to represent a given domain in an ontology to make data more sharable for software applications closely associated with that domain's needs, the less sharable the data becomes for other applications. For example, commitments made to develop a gene ontology useful for genomics applications render the ontology to be less likely to be used to share medical records data. Yet data sharing between genomics and patient care applications is critical in many situations. The remedy to this seeming conundrum is to identify general concepts that cut across many domains—the domains' "common-sense"—and collect them in a common-sense ontology (Lenat, 1995; Milton & Kazmierczak, 2004). This ontology is separated from more domain-specific ones. In fact, terms in several domain-specific ontologies can be defined using terms from a common general ontology. For example, a molecular biological ontology may provide building block representations for a biomedical ontology, which in turn underpins both gene and health care ontologies. The general ontologies also underlie an ontology of a different perspective—that of costing (Fox & Gruninger, 1998).

Arguably the ontologies of *emerging domains*, to be discussed below, make ontological commitments to their respective domains. It is prudent to ask, however, what are the common-sense ontologies that underlie these emerging domains?

Related Content

UB2SQL: A Tool for Building Database Applications Using UML and B Formal Method
Amel Mammarand Régine Laleau (2009). *Database Technologies: Concepts, Methodologies, Tools, and Applications (pp. 1168-1188).*
www.irma-international.org/chapter/ub2sql-tool-building-database-applications/7964

Agile Software Development in Practice
Matti Rossi, Hilkka Merisalo-Rantanenand Tuure Tuunanen (2007). *Research Issues in Systems Analysis and Design, Databases and Software Development (pp. 1-32).*
www.irma-international.org/chapter/agile-software-development-practice/28431

On the Representation of Temporal Dynamics
Salvatore T. Marchand Gove N. Allen (2003). *Advanced Topics in Database Research, Volume 2 (pp. 37-53).*
www.irma-international.org/chapter/representation-temporal-dynamics/4340

Evaluating Re-Identification Risks of Data Protected by Additive Data Perturbation
Han Li, Krishnamurty Muralidhar, Rathindra Sarathyand Xin (Robert) Luo (2014). *Journal of Database Management (pp. 52-74).*
www.irma-international.org/article/evaluating-re-identification-risks-of-data-protected-by-additive-data-perturbation/117744

Countermeasures for Protecting Legally Sensitive Web-Powered Databases and Web Portals
Theodoros Evdoridisand Theodoros Tzouramanis (2009). *Database Technologies: Concepts, Methodologies, Tools, and Applications (pp. 2079-2084).*
www.irma-international.org/chapter/countermeasures-protecting-legally-sensitive-web/8021