IDEA GROUP PUBLISHING
701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com
ITB12833

This paper appears in the book, *Emerging Trends and Challenges in Information Technology Management, Volume 1 and Volume 2* edited by Mehdi Khosrow-Pour © 2006, Idea Group Inc.

Human-Centric Challenges in Ontology Engineering for the Semantic Web: A Perspective from Patterns Ontology

Pankaj Kamthan & Hsueh-Ieng Pai
Dept of Computer Science & Software Engineering, Concordia University, Montreal, Quebec, Canada H3G 1M8,
T: (514) 848-2424-3000, F: (514) 848-2424-2830, {kamthan, hsueh_pa@cse.concordia.ca}

ABSTRACT

Ontologies form one of the most important layers in the Semantic Web architecture. We consider the symbiosis of human factors and ontology engineering, based on our experience in developing a large-scale ontology in the domain of Web Application patterns.

1. INTRODUCTION

The Semantic Web has recently emerged as an extension of the current Web that adds technological infrastructure for better knowledge representation, interpretation, and reasoning [8]. The development of domain-specific ontologies in the Web Ontology Language (OWL) [2] is critical to participation in the Semantic Web.

One of the goals of the Semantic Web is to facilitate "computers and people working in cooperation." However, neither the current Semantic Web architecture, nor the methodologies of ontology engineering [5] explicitly take into account or provide support for the human aspect of ontological development. From a semiotic viewpoint, the focus has primarily been on the *technical* side, that is, on syntactic and semantic levels, not on the pragmatic or social levels. It has been reported that ignoring the human dimension can potentially lead to ontologies that are irrelevant to the user [3] or have large variations in their semiotic quality [1].

The ultimate aim of our research is to construct a user-oriented, quality-driven, and descriptive markup-centric framework for the representation of software patterns. In this paper, we present a retrospective of the interplay between human factors and ontology engineering based on our experience in developing OWAP, an OWL ontology for typical structural patterns in a Web Application.

The rest of the paper is organized as follows. In Section 2, we give a brief outline of OWAP and its potential uses. Section 3 discusses the challenges of ontological representation of patterns from different perspectives in a human-centric ontology engineering project. Finally, Section 4 presents concluding remarks.

2. SYNOPSIS OF AN ONTOLOGY FOR WEB APPLICATION PATTERNS

Patterns are abstractions of knowledge gained from past experience and expertise in solving recurring problems at all levels of development [15]. In the past decade, patterns have been discovered in a variety of domains, including Web Applications [16], where they have proven to be useful instruments of guidance and reference to Web Engineers.

The growing number of patterns calls for their effective management [11]. The Semantic Web provides a suitable vehicle for communication and ontologies serve as an appropriate medium for representing knowledge inherent in patterns [10]. This motivated us to engineer OWAP [12].

The process for engineering OWAP was human-centric: it was driven by utility goals, documented comprehensively, and inspected/tested frequently for quality attributes of concern to its users. The OWAP conceptual model is briefly described as follows. The Web Application patterns are divided into two categories. The first category consists of patterns that describe the possible components that a Web Application can *physically* be composed of. Now, a Web Application will normally not consist of all or an arbitrary combination of these patterns. Therefore, the second category consists of patterns that describe how the patterns can be organized *logically* so that a Web Application can be formed using patterns that make sense. These categories are modeled as class hierarchies in OWL. Now, each individual pattern has its unique defining properties and, as a collective, patterns are related to each other in some manner (such as via inheritance, composition, and so forth). These aspects are modeled using object and datatype properties in OWL.

OWAP enabled us to make "interesting" inferences, including derivation of facts not literally present in the ontology but entailed by the semantics, and answer certain commonly asked questions by users such as: *What kind of components is an E-Commerce Web Site composed of? In what situation (context) do I need to include a Privacy Policy Page in my Web Site? What patterns are related to the Travel Web Site pattern?*

However, the OWAP effort also exposed serious issues in the current ontology engineering environment that we believe needs to be communicated and addressed for the benefit of those who may plan for a similar undertaking.

3. HUMAN-CENTRIC CHALLENGES IN ONTOLOGY ENGINEERING OF WEB APPLICATION PATTERNS

The challenges presented in this section are non-mutually exclusive and inspired by notions from traditional software quality, interaction design, and cognitive psychology.

3.1. Domain-Specific Challenges

Due to their virtual nature, software patterns pose unique knowledge representational challenges as compared to other, more tangible, domains. Even if certain concept or relationship is crucial to the domain, it may not be possible to represent it (adequately, or at all) within the current ontology engineering environment.

The modality of information in the patterns domain presents the foremost representational challenge. Furthermore, there are currently no ontology specification languages for the Semantic Web that provide adequate support for representing aesthetical, spatial/temporal information, or for satisfactorily representing vagueness/uncertainty in knowledge. Therefore, for example, it is difficult to specify in an ontology that a push button as an image should not appear "flat", or that the items in a menu should be "close" to each other, or that a Web Page instance may link to another Web Page only "occasionally."

3.2. Collaboration Challenges

In general, domain-specific ontologies, such as OWAP, need to reflect consensus among people for their existence, sustenance, and long-term usefulness. In particular, ontologies for the domain of patterns (and for other similar domains) require *agreement* among people about concepts and relations among them. For practical purposes, including time and communication constraints, this limits the number of stakeholders involved to a small group. This in turn restricts the amount of available skills and expertise required for knowledge acquisition and modeling.

Furthermore, when working in a team, the interaction among stakeholders depends upon characteristics of human behavior [9]. These personality types [13] vary with respect to the *role* (such as knowledge acquisitor, designer, and tester) that is taken upon by the stakeholder in the ontology production. A successful completion of the ontology then intimately depends on the compatibility among the stakeholders.

3.3. Learnability Challenges

OWAP was implemented in OWL as it is the successor of a number of initiatives for ontology specification languages for the Semantic Web, due to its foundations in well understood declarative semantics, and due to available tool support.

However, OWL could be considered as a language with a steep learning curve: its language definition is not self-contained and (recursively) spread over several specifications, it has a relatively large number of constructs, and it consists of three sub-languages (OWL Lite, OWL DL, and OWL Full). Specifically, OWAP was implemented using OWL DL, which requires the user to have an in-depth understanding of the Description Logics formalism.

3.4. Authoring Challenges

An OWL ontology could be expressed in a variety of syntaxes, of which only Resource Description Framework (RDF)/Extensible Markup Language (XML) based serialization is normative. However, the RDF/XML syntax of even a single concept can be prohibitively verbose for human use and error-prone particularly due to the complexity of relationships involved. This imposes the necessity for a non-textual (and, say, graphical) approach to authoring.

Protégé-2000 [14] was used as the primary ontology authoring environment for OWAP. However, we had to overcome certain idiosyncrasies that it presented: the software was at times resistant to certain necessary modifications in the ontology, its interface had refresh issues, and it did not preserve the structure of markup of a file created using other tools. The need for an improved cognitive support in ontology tools such as Protégé-2000 has been emphasized [4].

We note that the imposition of a graphical approach to authoring evidently presents an obstacle towards participation by those with some form of visual disability. This is not in alignment with the vision of inclusiveness of the Semantic Web for all users.

3.5. Reasoning Challenges

The reasoner used to derive inferences from OWAP was Racer [7], which is a complete and fairly stable support for OWL DL. It was able to answer most of the queries within a few seconds. However, as the number of instances in OWAP becomes large, performance can become an issue. For example, a response for a query that involves the use of a transitive property can take a few hours. This is prohibitive for practical purposes in a decentralized environment such as the Semantic Web.

Reasoning with OWAP may not be readily possible on certain devices. For example, the query formulation in Racer for the question *What are the known uses of the Hot List pattern?* is:

```
(individual-fillers |http://a.com/owap#WebPageComponent_HotList|
|http://a.com/owap#hasKnownUse|)
```

This turns out to be prohibitively lengthy for use in environments with restricted interface capabilities such as a cellular phone or a Personal Digital Assistant (PDA) client.

4. CONCLUSION

The Semantic Web holds much promise as a platform for a new breed of applications that by virtue of increased machine automation would reduce some of the tedium involved in human processing. Therefore, we need to recognize that the purpose of the Semantic Web is to serve the needs of humans, not that of the machines, and work to that effect.

The development of an ontology such as OWAP is a result of both individual activity and a collaborative social process. The mental and physical capabilities and preferences of an individual, or the manifestations of these in a group of individuals, play a central role in ontology engineering. Recognizing the human dimension of ontology engineering can help providers plan appropriately and make user experience more predictable.

As a future research avenue, we plan to evaluate OWL using the Cognitive Dimensions of Notations [6], which is a generic framework for describing the utility of information artifacts by taking the system environment and the user characteristics into consideration. We intend to use the results as a basis for a systematic approach to quality assurance and evaluation of ontological representations of patterns.

REFERENCES

1. Burton-Jones, A., Storey, V. C., Sugumaran, V., Ahluwalia, P. (2005). A Semiotic Metrics Suite for Assessing the Quality of Ontologies. Data & Knowledge Engineering, 55(1), 84-102.
2. Dean, M., Schreiber, G. (2004). OWL Web Ontology Language Reference. W3C Recommendation. World Wide Web Consortium (W3C).
3. Degler D., Lewis R. (2004). Maintaining Ontology Implementations: The Value of Listening. Proceedings of the Extreme Markup Languages 2004 Conference, Montreal, Canada, August 2-6, 2004.
4. Ernst, N. A., Storey, M.-A., Allen, P. (2005). Cognitive Support for Ontology Modeling. International Journal of Human-Computer Studies, 62(5), 553-577.
5. Gómez-Pérez, A., Fernández-López, M., Corcho, O. (2004). Ontological Engineering. Springer Verlag.
6. Green, T. R. G. (1989). Cognitive Dimensions of Notations. In: Sutcliffe, V. A., Macaulay, L. (Editors). People and Computers. Cambridge University Press. 443-460.
7. Haarslev, V., Möller, R. (2001). Description of the Racer System and its Applications. Proceedings of the 2001 International Workshop on Description Logics (DL 2001), Stanford, USA, August 1-3, 2001, 132-141.
8. Hendler, J., Lassila, O., Berners-Lee, T. (2001). The Semantic Web. Scientific American, 284(5), 34-43.
9. Jung, C. G. (1971). Psychological Types. Princeton University Press.
10. Kamthan, P. (2005). A Framework for Representation of Software Patterns. Proceedings of the IBIMA 2005 Conference on Theory and Practice of Software Engineering for the 21st Century (TPSE 2005), Cairo, Egypt, December 13-15, 2005.
11. Kamthan, P., Pai, H.-I. (2006). Knowledge Representation in Pattern Management. In: Encyclopedia of Knowledge Management. Schwartz, D. (Editor). Idea Group Publishing.
12. Kamthan, P., Pai, H.-I. (2006). Representation of Web Application Patterns in OWL. In: Web Semantics and Ontology. Taniar, D., Rahayu, J. W. (Editors). Idea Group Publishing.
13. Keirsey, D. (1998). Please Understand Me II. Prometheus Nemesis Book Company.
14. Noy, N. F., Sintek, M., Decker, S., Crubezy, M., Fergerson, R. W., Musen, M. A. (2001). Creating Semantic Web Contents with Protege-2000. IEEE Intelligent Systems, 16(2), 2001, 60-71.
15. Rising, L. (1999). Patterns: A Way to Reuse Expertise. IEEE Communications Magazine, 37(4), 34-36.
16. Van Duyne, D. K., Landay, J., Hong, J. I. (2003). The Design of Sites: Patterns, Principles, and Processes for Crafting a Customer-Centered Web Experience. Addison-Wesley.