Editorial

Where is the Web in the Semantic Web?

This special issue collects together selected papers from the Semantic Web track of the 2006 World Wide Web Conference. The central idea of the Semantic Web is to extend the current human-readable Web by encoding some of the semantics of Web resources in a machine-processable form. Moving beyond syntax opens the door to more advanced applications and functionality on the Web. Computers are better able to search, process, integrate and present the content of these Web resources in a meaningful, intelligent manner.

In prior semantic Web tracks of the WWW conference, much good work was reported on the semantic side of the “Semantic Web”. Contributions came from many different fields such as databases, natural language processing, machine learning, information retrieval, knowledge representation, and others. This important work has laid the foundation for the Semantic Web.

Our goal for the Semantic Web track of the 2006 WWW Conference was to re-emphasize the Web aspects of the Semantic Web. We wanted to better understand how semantics can provide new levels of Web functionality, either for end users, or for designers and developers. The emphasis was on practical aspects, rather than theoretical ones. We asked for papers that would advance our understanding of how semantic technologies can be exploited on the Web. We explicitly called for papers that:

• show how semantic technologies add value to the Web, achieving things that alternative technologies cannot do as well, or at all;
• present new semantic technologies, or novel applications of existing semantic technologies that provide new levels of Web functionality;
• present new Web technologies, or novel applications of existing Web technologies that, when combined with semantic technologies add new Web functionality.

We are pleased to present four papers selected from that track, revised for this special issue. They cover the gamut from some pioneering work in the area of trust in Web content, to a comparison of two leading approaches for representing knowledge on the Web, to a semantic enhancement to the infrastructure underpinning the Wikipedia, to an advanced social network extraction system which has been used at four academic conferences.

Towards Content Trust of Web Resources—Jolanda Gil and Donovan Artz describe some new work in the area of trust in content on the Semantic Web—an area that has seen little attention to date. Most prior work on trust focuses on issues such as authentication and reputation, and does not take into account the nature and use of the information itself. They describe an interesting and detailed study identifying and analyzing just what factors people use to decide what content they will trust and why. The goal is to be able to capture as much of this information as possible semi-automatically with minimal user interaction. A simulation environment is described to study alternative models of content trust.

A Comparison of Two Modelling Paradigms in the Semantic Web—Ian Horrocks and Peter Patel-Schneider, two well-known theorists, consider the differences between classical logics and datalog-related logics, assessing their suitability for practical Semantic Web applications. They raise a wide range of technical and practical issues and conclude that the open environment of the Semantic Web is better served by standard logics. Although firmly in one camp, the authors do an admirable job of presenting both sides of the issues. Of course, for those in the other camp, it may seem less balanced.

Semantic Wikipedia—Max Volkel, Markus Kroetzch, Denny Vrandecic, Heiko Haller and Rudi Studer present in detail the motivation and design of a semantic extension to the code base underlying the Wikipedia, the world’s largest collaboratively edited source of encyclopaedic knowledge. They show how using semantic annotations significantly enrich the content, so that it can be browsed, searched, and reused in novel ways. The first three papers make important strides that should facilitate and impact our ability to build future practical Semantic Web applications. Our final paper describes an actual application, along with some novel techniques.

POLYPHONET: An Advanced Social Network Extraction System from the Web—Yutaka Matsuo, Junichiro Mori and Masahiro Hamasaki . . .

The paper describes a social network extraction system called POLYPHONET, which employs several advanced techniques to extract relations of persons, detect groups of persons, and obtain keywords for a person. Traditional search engines are used to measure co-occurrence of information and obtain Web documents. POLYPHONET advances beyond existing search
engines to extract social networks by classifying social relations into categories, and by obtaining and utilizing person-to-word relations. The system has been in use at four academic conferences, each with more than 500 participants.

Together, these four papers cover a broad spectrum from theory to application, and from first steps to solid engineering. They are all concerned with the combination of semantic techniques with the Web as it exists “out there in the real world”. For these reasons, we have highly recommended these papers.

Frank van Harmelen
Michael Uschold *
Phantom Works, M&CT, P.O. Box 3707, Seattle, WA 98124, United States

* Corresponding author. Tel.: +1 425 373 2845.
E-mail address: michael.f.uschold@boeing.com (M. Uschold)

Available online 15 September 2007