

Semantic Web, Database Management and Information Systems: Overview of the Special Issue

Amicalola workshop on DB-IS Research for Semantic Web and Enterprises was held in April 2002 (<http://lsdis.cs.uga.edu/SemNSF/>). Secluded and cozy retreat in Georgia mountains provided an excellent environment for high quality discussions between a set of invited senior researchers from within and outside of database and information systems community. We are delighted to present the final report from the workshop, which attempts to capture the discussions and conclusions of the workshop participants (Amit Sheth and Robert Meersman: Amicalola Report: Database and Information Systems Research Challenges and Opportunities in Semantic Web and Enterprises).

Building upon the momentum of this workshop, we are also delighted to present a set of invited papers in this special issue of SIGMOD Record. As Semantic Web gains momentum, this is an attempt to further explore and expose the thinking and work of database and information systems researchers targeted to enabling the Semantic Web. Ten papers can be divided into three general areas:

Methodologies, Frameworks and Architectures

In the first part of their paper, Spyns, Meersman and Jarrar (Data Modelling versus Ontology Engineering) discuss the differences between data modeling vs. ontology engineering. They specify four differentiating criteria that help to understand the genericity of ontologies (i.e. their relative independence from particular applications) versus the relative specificity of data models (e.g. database schemas). In the second part, the authors illustrate this in an ontology engineering approach -called DOGMA- that supports and favors the genericity of ontologies. This approach separates "atomic" conceptual relations from "predicative" domain rules. Under this fundamental separation, an ontology in DOGMA consists of an "ontology base" that holds sets of intuitive context-specific conceptual relations and a layer of "ontological commitments" which hold the domain rules. By doing so, the authors claim a maximization of knowledge reusability and achieving scalability in ontology modeling.

Sure, Staab and Studer (Methodology for Development and Employment of Ontology based Knowledge Management Applications) illustrate a methodology for introducing and maintaining ontology based knowledge management applications into enterprises with a focus on Knowledge Processes and Knowledge Meta Processes. While the former process circles around the usage of ontologies, the latter process guides their initial set up. The methodology is illustrated by an example from a case study on skills management.

Bussler, Fensel and Maedche (A Conceptual Architecture for Semantic Web Enabled Web Services) state that Web Services in their current state are insufficient to address real semantic interoperability. It is argued that enriching Web Services with Semantic Web technologies like ontology support will result in real interoperable web services termed Semantic Web enabled Web Services (SWWS). The Web Service Modeling Framework (WSMF) is introduced that defines the necessary concepts for SWWS providing maximal de-coupling, scalable mediation and support for behavior level. An architecture implementing the components necessary to deploy SWWS is introduced.

Finin and Joshi (Agents, Trust, and Information Access on the Semantic Web) deal with some key issues in their effort to transform the web to an active ecology of agents that produce, consume and act on the information. The first provide an overview of work in dynamic composition of data and services and in indexing the hidden web. More importantly, they introduce challenges related to distributed trust and belief, a critical challenge that must be addressed to enable broader Semantic Web adoption. In particular, they emphasize the need for two main components in order to secure the Semantic Web: a semantic policy language for defining security requirements and a distributed trust management approach.

Maximilien and Singh (A Conceptual Model of Service Reputation) present an approach that provides a conceptual model for reputation that captures the semantics of attributes. The semantics includes characteristics, which

describe how a given attribute contributes to the overall rating of a service provider and how its contribution decays over time. Maximilien and Singh's approach thus applies both to reputations and to explicit endorsements of a service provider by another party. In Finin and Joshi's proposal, a principal might trust an object if that object is trusted by a third party that is trusted by the given principal. This is similar to Maximilien and Singh's notion of endorsement. A key difference between the two approaches is that Finin and Joshi seek to capture policies for endorsement and delegation, whereas Maximilien and Singh seek to capture service attributes and how they can be combined to support various policies.

Concepts and Techniques

Anyanwu and Sheth (The ρ Operator: Discovering and Ranking Semantic Associations on the Semantic Web) emphasize that relationships, more than entities, are at the heart of representing semantics. Identification, discovery, validation and utilization of relationships will be a critical computation on the Semantic Web. They introduce an operator for expressing queries about complex relationships (semantic associations) that uses a notion of context, which allows for restricting the search space and for context-driven ranking of results. Just as a Web search engine looks for relevant documents in the current Web, this operator ρ can be seen as discovering and ranking complex relationships in the Semantic Web.

Aberer, Cudré-Mauroux and Hauswirth (A Framework for Semantic Gossiping) identify lack of semantic interoperability as the key impediment in broadening the impact of P2P systems, and at the same time view the emerging P2P paradigm as an opportunity for semantic interoperability. Most notable aspect of this paper is a proposal for a distributed strategy of achieving semantic agreement, which is relevant to achieving higher scale in the Semantic Web. Their approach called Semantic Gossiping allows users to introduce their own schemas, but aims to achieve consensus or convergence by exchange of mappings or translations and using cycles for reaching intermediate closures.

Grosky, Sreenath, and Fotouhi (Emergent Semantics and the Multimedia Semantic Web) seek to complement fixed semantics through annotations (such as that by an author of content) by an emergent semantics defined through use of that document. They present an approach to capture dynamic context of content by analyzing users' browsing path.

Applications and Synergy

Several domains, including biology, national security, business intelligence and enterprise applications (such as CRM) provide vast playground for utilizing semantics in improving interoperability, computation and discovery. Buttler, Coleman, Critchlow, Fileto, Han, Liu, Pu, Rocco, and Xiong (Querying Multiple Bioinformatics Information Sources: Can Semantic Web Research Help?) give an insightful portrayal of what is involved in a biological discovery related to impact of environmental mutagen, showing where semantics can help. They then outline research challenges for DB/IS researchers in the areas of process management, resource discovery, data quality and data integration who can look into the promise of applying semantics to solve the challenging problems inspired by biology research.

The Grid is an emerging platform to support on-demand "virtual organisations" for coordinated resource sharing and problem solving on a global scale. Goble and De Roure (The Grid: An Application of the Semantic Web) discuss how the Grid is beginning to exploit technologies developed for Web Services and the Semantic Web. Conversely, the Grid and its scientific users provide application pull which will benefit the Semantic Web.

Furthermore, Grosky et al., and Aberer et al., discuss synergy and applications with multimedia and P2P infrastructure, respectively.

We also encourage the readers to review the discussions on what was termed as Semantic Web Information System during NSF PI's meeting. This, with various presentations and position papers are available at the Amicalola Workshop Web site.

Robert Meersman and Amit Sheth (Editors)