Introduction

Tharam S. Dillon¹, Elizabeth Chang¹, Robert Meersman², and Katia Sycara³

¹ Digital Ecosystems and Business Intelligence Institute
Curtin University of Technology

G.P.O. Box U1987, Perth, WA 6845, Australia
{Tharam.Dillon, Elizabeth.Chang}@cbs.curtin.edu.au

² STARLab, Vrije Universiteit Brussel, Belgium
meersman@vub.ac.be

³ The Robotics Institute, Carnegie Mellon University
katia@cs.cmu.edu

Abstract. The Web has now been in existence for quite some time and it is pervasive in its influence on all aspects of society and commerce. It has also produced a major shift in our thinking on the nature and scope of information processing. However in its technological nature and its supporting theoretical foundations, it was relatively rudimentary, being largely suitable for information dissemination. It is rapidly moving away from this, to application deployment and knowledge deployment that require complex interactions and properly structured underlying semantics. This has been a sudden upsurge of research activity in the problems associated with adding semantics to the Web. This work on semantics will involve data, knowledge, and process semantics.

1 Introduction

The emergence of Semantic Web (SW) and the related technologies promise to make the web a meaning full experience. Tim Berners-Lee¹ envisioned that the Semantic Web paradigm promises a meaningful web for both humans and the machines (agents). It provides representation, integration and access to information using semi-structured data models where data and meta-data semantics are represented in a way that is easily interpreted by autonomous agents. However, high level modeling, design and querying techniques prove to be challenging tasks for organizations that are hoping to utilize the SW paradigm for their commercial and industrial applications.

2 Themes

In order to help understand and interpret the research issues that are arising in this challenging area and the proposals for their solution, we have proposed a series of books, which is an extended journal format. The series will have periodic state of the art volumes and intermediate volumes directed towards hot button issues. The series will address a number of themes and these include but are not limited to:

¹ Berners-Lee, T. *Weaving the Web*. San Francisco: Harper, 1999.

Theme 1: Semantics and Concepts for Semantic Web

- Conceptual semantics
- Formal models
- Theoretical models for SW
- Semantics of agent and Web interaction
- Security and trust for the semantic Web
- Database technologies for the semantic Web
- Impact of semantic Web computing on organizations and society
- Semantics for ubiquitous computing
- Conceptual correctness in SW modeling

Theme 2: Emerging New Trends

- Query models and languages
- Ontology dynamics and evolution
- Ontology-Driven IS
- Human centered aspects for SW
- Trust & Security

Theme 3: Ontological Engineering

- Metadata and knowledge markup
- Ontology-Driven IS and Applications
- Ontological correctness in SW modeling
- New and emerging Standards
- Query models for knowledge extraction
- Formal models and language specification

Theme 4: Language for Semantic Web

- Extensions to languages
- Languages for knowledge extraction and integration
- Content-based information and knowledge retrieval
- Database technologies for the semantic Web
- Tools & Prototypes

Theme 5: Interoperability, Integration and Composition

- Concepts, Frameworks and Platforms
- Interoperability of data and Web services
- Ontology-Driven Interoperability and Integration for IS

Theme 6: Applications of Semantic Web in Domains

- Formal and Practical knowledge representation and inference for the semantic Web
- Human centered aspects specifically for the semantic Web
- Ontology-Driven IS Applications
- Health Information systems
- Bioinformatics

Theme 7: SOA and Web Semantics

- Web Services and SW
- Ontology bases and SOA

Theme 8: Industrial Applications & Experiences

- Prototypes
- Experience
- Security and trust for the semantic Web
- Multimodality and visualization technologies for the semantic Web
- Applications on mobile devices
- Application in Bioinformatics and Health domains

This is the first book in this series and is titled “Advances in Web Semantics: A State-of-the Art”. In addition, there will be one (or more) periodic theme driven book/(s) issues, one specific theme (such as semantics and concepts for semantic web, process semantics for the web, etc), that is of utmost importance to the advancement of the web semantics and its future.

3 Author Contributions

The contributions in this book are divided into three broad categories namely: Ontologies, Web Services and Applied Semantic Web. There are 6 chapters in Ontologies and Knowledge Sharing, 3 chapters in Web Services and 5 chapters in Applied Semantic Web respectively. Each chapter of the current book is carefully written and presented by selected authors and by the IFIP workgroup members, who are leaders in their field of expertise from wide ranging domains. The author selection process is by invitation only and the selection process is highly selective and should be approved by the editors and the IFIP workgroups. Each Chapter was then peer reviewed by at least two members of the working group and revised in accordance with the input from these reviews.

Chapter 2 presents a methodological framework for ontology engineering (called DOGMA), which helps in building ontologies that are highly reusable and usable, easier to manage and to maintain. The authors discuss the major challenges in ontology engineering. They also discuss ontology reusability verses ontology usability. Authors present the DOGMA approach, which prescribes that ontology, be built as separate domain axiomatization and application axiomatizations. While a domain axiomatization focuses on the characterization of the intended meaning of a vocabulary at the domain level, application axiomatizations focus on the usability of this vocabulary according to certain perspectives. In this chapter, the authors also show how ontology specification languages can be effectively (re)used in ontology engineering.

Process mining assists automatic discovery of information about process models in organizations. In chapter 3, the authors show how the analysis provided by current techniques can be improved by including semantic data in event logs. The chapter is divided into two main parts. The first part illustrates current process mining techniques and shows usage of the process mining tool ProM to answer concrete questions.

The second part utilizes usage scenarios to depict benefits of semantic annotated event logs and they define a concrete semantic log format for ProM.

Current large scale knowledge sharing methods have relied on pre-engineering of content. In chapter 4 the authors describe how to focus on semantics related to interactions. Their method is based on interaction models that are mobile in the sense that they may be transferred to other components, this being a mechanism for service composition. By shifting the emphasis to interaction knowledge sharing of sufficient quality for sustainable communities can be obtained without the barrier of complex meta-data provision.

Extracting high quality ontology from a given base ontology is needed in versioning, distribution and maintenance of ontologies. Chapter 5 presents a formalism one to extract the sub-ontology from a base ontology. In addition to extracting the sub-ontology it carries optimization of its structure using different optimization schemes. Authors discuss the overview of the formalism demonstrating several optimization schemes. Examples of how the formalism is deployed to reach a high-quality result, called a materialized ontology view, are covered in the chapter. The methodology provides a foundation for further developments, and shows the possibility of obtaining usable ontologies by automation.

Ontologies are one of the basic operations of semantic heterogeneity reconciliation. The aim of Chapter 6 is to motivate the need for ontology matching, introduce the basics of ontology matching, and then discuss several promising themes in the area. In particular, authors focus on such themes as uncertainty in ontology matching, matching ensembles, and matcher self-tuning. Finally, authors outline some important directions for future research.

In Chapter 7 the authors explore the development of systems for software engineering ontology instantiations management in the methodology for multisite distributed software development. In a multisite distributed development environment, team members in the software engineering projects must naturally interact with each other and share lots of project data/agreements amongst themselves. Since they do not always reside at the same place and face-to-face meetings are expensive and very infrequent, there is a need for a methodology and tools that facilitate effective communication for efficient collaboration. Whilst multi-site distributed teams collaborate, there are a lot of shared project data updated or created that leads to a large volume of project information and issues. Systematic management of Software engineering knowledge is represented in the software engineering ontology.

The Web infrastructure and the availability of Semantic Web-style metadata have enabled new types of inter-organizational business processes, which in turn need new business process models. Current approaches focus on producing taxonomies or categorizations, or on stating what aspects of business process models should be included in (or excluded from) modeling. In Chapter 8, authors discussed the different facets such a representation must possess, starting from standard ones like workflow representation and arriving to the capability of defining the structure of a company's value chain and describing the position of the process actors within the value network. While authors believe a visual approach to modeling is preferable from the modeler's point of view, they focused on the logics-based models underlying each facet of the process representation. They build on the idea that the different facets of a business process model require different formalizations, involving different inference techniques,

some of them Semantic Web-style, others closer to classical Prolog reasoning. Also using a simple business process model, they discussed how well our multi-faceted representation can be integrated by a hybrid approach to reasoning, transferring knowledge and metrics obtained reasoning on each part of the model to the other parts.

In Chapter 9, the authors look at some of the issues associated with the problem of privacy protection, and present some directions of work that would help automate the solution to the problem. They have also presented architecture for privacy protection at the back end that comprises of simple rules that specify which user items can access which data items or categories.

Current search engines that are based on syntactic keywords are fundamental for retrieving information often leaving the search results meaningless. Increasing semantics of web content would solve this problem and enable the web to reach its full potential. This would allow machines to access web content, understand it, retrieve and process the data automatically rather than only display it. In chapter 10 authors illustrate how the ontology and multi-agent system technologies, which underpin this vision, can be effectively implemented within the mental health domain.

The manufacturing sector is currently under pressure to swiftly accommodate new products by quickly setting up new factories or retrofitting existing ones. In order to achieve this goal, engineering tasks currently performed manually need to be automated by using ontologies. Chapter 11 reviews current approaches to use ontologies in the manufacturing domain, which include use for vocabulary definition in multi-agent systems and use for describing processes using Semantic Web Services. In addition, current and emerging research trends are identified.

Before initiating a financial e-commerce interaction over the World Wide Web, the initiating agent would like to analyze the possible Risk in interacting with an agent, to ascertain the level to which it will not achieve its desired outcomes in the interaction. To determine the possible risk in an interaction, the initiating agent has to determine the probability of failure and the possible consequences of failure to its resources involved in the interaction. In Chapter 12 as a step towards risk analysis, authors propose a methodology by which the initiating agent can determine beforehand the probability of failure in interacting with an agent, to achieve its desired outcomes.

The ability to deal with incompatibilities of service requesters and providers is a critical factor for achieving smooth interoperability in dynamic environments. Achieving interoperability of existing web services is a complicated process including a lot of development and integration effort, which is far from being automated. Semantic Web Services frameworks strive to facilitate flexible dynamic web services discovery, invocation and composition and to support automation of these processes. In chapter 13, authors focus on mediation and brokering mechanisms of OWL-S Web Services, which overcomes various types of problems and incompatibilities.

The recent proliferation of SOA applications exacerbates this issue by allowing loosely-coupled services to dynamically collaborate with each other, each of which might maintain a different set of ontologies. Chapter 14 presents the fundamental mechanism of Latent Semantic Analysis (LSA), an extended vector space model for Information Retrieval (IR), and its application in semantic web services discovery, selection, and aggregation for digital ecosystems. Firstly, the authors explore the nature of current semantic web services within the principle of ubiquity and simplicity. This is followed by a succinct literature overview of current approaches for

semantic services/software component discovery and the motivation for introducing LSA into the user-driven scenarios for service discovery and aggregation. They then direct the readers to the mathematical foundation of LSA – SVD of data matrices for calculating statistics distribution and thus capturing the ‘hidden’ semantics of web services concepts. Some existing applications of LSA in various research fields are briefly presented, which gives rise to the analysis of the uniqueness (i.e. strength, limitations, parameter settings) of LSA application in semantic web services. They also provide a conceptual level solution with a proof-of-concept prototype to address such uniqueness. Finally they propose an LSA-enabled semantic web services architecture fostering service discovery, selection, and aggregation in a digital ecosystem.

Web services have been emerging as the lead implementation of SOA upon the Web, adding a new level of functionality for services description, publication, discovery, composition and coordination. We still need to combine web services in a useful manner, which limits scalability. A common agreement is the need to semantically enrich Web Services description. Semantic web services have been emerging with promising potential to satisfy SOA requirements. Chapter 15 is divided into three parts. Firstly, authors depict the concepts and principles of SOA. They highlight the dynamicity, flexibility and the challenges it poses. In the second part, the authors present Web services as the key technology implementing SOA principles. They state in particular the main standards and technologies and investigate how they fit in SOA and how far they satisfy its requirements. Finally, authors exhibit SWS conception and foundation. They consider the two SWS languages WSMO and OWL-S show how these promise to overcome SOA challenges.