Editorial: Revisiting the (Machine) Semantic Web: The Missing Layers for the Human Semantic Web

Gottfried Vossen, Miltiadis Lytras, and Nick Koudas

THE Semantic Web has recently emerged as a new and highly promising context for knowledge and data engineering. Within an atmosphere of high expectations, many myths as well as many visions have exhibited a number of different approaches for the exploitation of the Semantic Web in both academia and industry. However, a struggling business reality requires a concrete strategy as well as the development of specific competencies from the knowledge and data engineering community in order to prove the value of the Semantic Web to society.

For three years, we have undertaken a significant effort to cultivate the Semantic Web vision in the computer science, information systems, and Semantic Web communities through the Special Interest Group on Semantic Web and Information Systems of the Association for Information Systems (AIS, see http://www.aisnet.org [1], [8]). We have emphasized the benefits of the Semantic Web merits in different application contexts, including digital libraries, e-government, knowledge management, health care, and e-learning [3], [4], [5], [6]. One of the most fascinating aspects of our effort is the exchange of ideas with people that come from different disciplines, which has led us to arrive at the conclusion that the Semantic Web adoption requires the convergence of many different disciplines; this is illustrated in Fig. 1.

It is more than obvious that, after an early stage of evolution, Semantic Web research has reached a first level of maturity. Most significantly, some "voices" of criticism or questioning for the pace of the change that Semantic Web brings to traditional knowledge and data engineering have initiated a new stream of innovations. The concept of semantics [2] and its capacity to support a new era of applications challenges the traditional perceptions for the never-ending journey of computing. Knowledge and data representation and retrieval require new conceptual models and the move to a human Semantic Web vision seems more timely than ever. After the initial enthusiasm and excitement following the launch of the Semantic Web vision, a time of significant problems, unexploited opportunities, and slow adoption followed. Knowledge and data engineering had to meet several diverse and high demanding requirements for the realization of the Semantic Web. In numerous international efforts as well as in various research and competence centers of the Semantic Web, there is now a continuous effort to reach the point for a real "take-off" of the Semantic Web. It looks like an entire research community is looking for the last step before the breaking of the "wall." And, this wall is associated with all the inefficiencies of the traditional Web, and with a panacea for the solution of all the knowledge-related performance gaps.

For the past one and a half years, we have been working hard, with the support of a large number of reviewers, on the development of an excellent quality *TKDE* special issue on the Semantic Web. From the beginning, our motivation was based on a clear belief that the Semantic Web represents a key milestone for the knowledge and data engineering community. While the Semantic Web is often considered to be a machine-intensive/oriented theme, our key argumentation is that Semantic Web is a Human manifesto [7]. The fundamental social and political impact of the Semantic Web is derived from the fact that its underlying technology supports a shift of social interaction patterns from "knowledge push" to "knowledge pull" [7]. This includes the shift

- from teacher-centric to learner-centric education,
- from doctor-centric to patient-centric health care,
- from bureaucrat-centric to citizen-centric administration,
- from government-centric to citizen-centric democracy, and
- from producer-centric to consumer-centric business models.

The technical reason for enabling such shifts is not that machines can understand what we talk about, something they will, of course, never be able to do, but the fact that on the Semantic Web, machines are able to decide whether or not we are talking about the same thing. Hence, a more semantically correct name for the Semantic Web would be something like the "Identity Conflict Resolvable Web." This property of the Semantic Web enables the information about the information (i.e., the metadata) to become as distributed as the information itself. This makes it possible to move from "opinion registration," which is the basis of knowledge push, to "opinion publication," which is the

- *G. Vossen is with the European Research Center for Information Systems (ERCIS), University of Muenster, Leonardo-Campus 3, 48149 Muenster, Germany. E-mail: vossen@helios.uni-muenster.de.*
- *M. Lytras is with the Computer Engineering and Informatics Department, University of Patras, Argolidos 40-42, 153 44 Gerakas Attikis, Greece. E-mail: lytras@ceid.upatras.gr.*
- *N. Koudas is with the Department of Computer Science, Bahen Center for Information Technology, University of Toronto, 40 St. George Street, Rm BA5240, Toronto ON M5S 2E4. E-mail: koudas@cs.toronto.edu.*

For information on obtaining reprints of this article, please send e-mail to: tkde@computer.org.

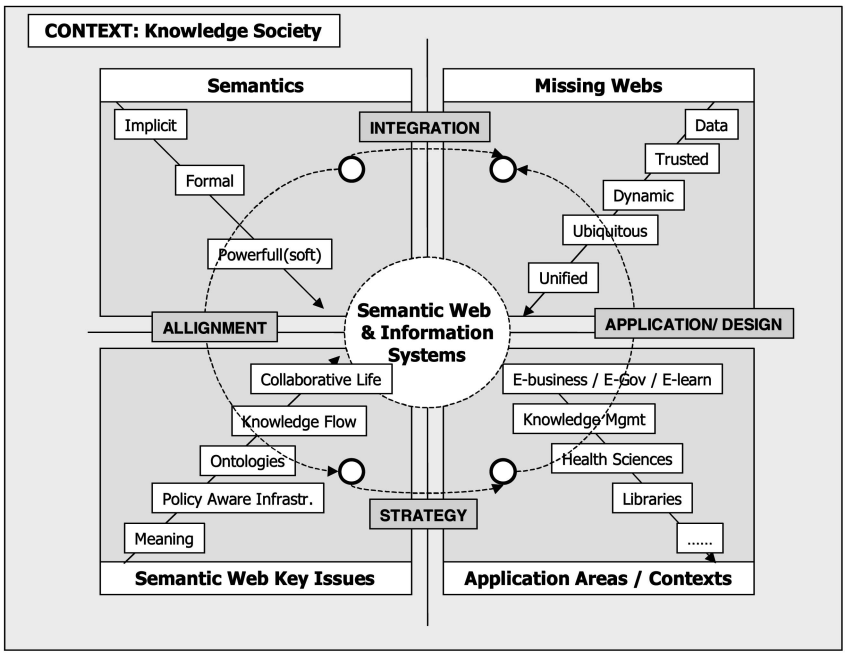


Fig. 1. Illustration of Semantic Web adoption.

basis of knowledge pull. This is the "root cause" from which the social and political power of the Semantic Web originates.

Our special issue has a three-fold integrated contribution:

1. It **delivers the state-of-the-art** in the current Semantic Web research worldwide, by communicating excellent research papers.
2. It **promotes the discipline** by synthesizing different perspectives, by integrating complimentary pieces that contribute to the-state-of-the-art, and by innovating through sound propositions.
3. It **challenges the future research** by indicating open research issues that will enhance the Semantic Web adoption and capacity to address key problems.

The high-quality publication standards of the *IEEE TKDE* were taken into consideration in the total spectrum of the development process of the Semantic Web special issue. The final acceptance of 14 research papers out of 60 submissions, after three very intensive and focused rounds of blind review and with the support of almost 200 capable reviewers, concludes an intellectual task that we have carried out with passion and inspiration. The accepted papers in this special issue fall into two categories:

1. **Basic Tools and Languages**
   a. *Metadata/Taxonomies Extraction and Refinement:*
      i. "Bottom-Up Extraction and Trust-Based Refinement of Ontology Metadata," by Paolo Ceravolo, Ernesto Damiani, and Marco Viviani, presents a way of building ontologies that proceed in a bottom-up fashion, defining concepts as clusters of concrete XML objects. Their key contribution substantially improves metadata quality in the long run.
      ii. "Mining Generalized Associations of Semantic Relations from Textual Web Content," by Tao Jiang, Ah-Hwee Tan, and Ke Wang, presents a two-step procedure to mine generalized associations of semantic relations conveyed by the textual content of Web documents. Their key contribution is a novel generalized association pattern-mining algorithm (GP-Close).
      iii. "A Taxonomy Learning Method and Its Application to Characterize a Scientific Web Community," by Paola Velardi, Alessandro Cucchiarelli, and Michaël Pétit, presents a semi-automated strategy to extract domain-specific taxonomies from Web documents and its application to model a Network of Excellence in the emerging research field of enterprise interoperability.
   b. *Semantic Interoperability:*
      i. "RDFS(FA): Connecting RDF(S) and OWL DL," by Jeff Z. Pan and Ian Horrocks, proposes a novel modification of RDF(S) as a firm semantic foundation for many of the latest Description Logics-based SW ontology languages, including OWL DL. As a result, the introduction of RDFS(FA) solidifies RDF(S)'s proposed role as the base of the Semantic Web and facilitates key knowledge engineering tasks.
      ii. "Introducing Time into RDF," by Claudio Gutierrez, Carlos A. Hurtado, and Alejandro Vaisman, presents a framework to incorporate temporal reasoning into RDF, yielding temporal RDF graphs, and discusses in detail various aspects of their

novel approach including semantics, timestamps, and a temporal query language for RDF.

iii. "Interoperability Support between MPEG-7/21 and OWL in DS-MIRF," by Chrisa Tsinaraki, Panagiotis Polydoros, and Stavros Christodoulakis, describes the DS-MIRF Framework, a software engineering framework that facilitates the development of knowledge-based multimedia applications such as multimedia information retrieval, filtering, browsing, interaction, knowledge extraction, segmentation, and content description.

c. *Reasoning:*

i. "DR-Prolog: A System for Defeasible Reasoning with Rules and Ontologies on the Semantic Web," by Grigoris Antoniou and Antonis Bikakis, reports on the implementation of a system for defeasible reasoning on the Web that is syntactically compatible with RuleML and has many novel characteristics.

ii. "A Flexible Ontology Reasoning Architecture for the Semantic Web," by Jeff Z. Pan, proposes a flexible reasoning architecture for OWL-Eu and OWL-E, two decidable extensions of the W3C standard ontology language OWL DL and describes their prototype implementation with very interesting characteristics.

2. **Applications & Contexts of Exploitation**

a. *Search Engines—Novel Approaches:*

i. "An Adaptation of the Vector-Space Model for Ontology-Based Information Retrieval," by Pablo Castells, Miriam Fernández, and David Vallet, proposes a model for the exploitation of ontology-based knowledge bases to improve search over large document repositories. It contributes to the problem of Semantic Search by introducing a search engine that returns documents rather than, or in addition to, exact values in response to user queries.

ii. "A Relation-Based Search Engine In Semantic Web," by Yufei Li, Yuan Wang, and Xiaotao Huang, proposes a prototype relation-based search engine, "OntoLook," which has been implemented in a virtual Semantic Web environment and exploit an interesting algorithm.

b. *Grid and Knowledge Management:*

i. "A Semantic Web-Based Approach to Knowledge Management for Grid Applications," by Liming Chen, Nigel R. Shadbolt, and Carole A. Goble, analyzes the nature of Grid computing and its requirements for knowledge support. A Semantic Web-based approach is proposed in the context of a real-world Grid application, the GEODISE project.

ii. "A Requirements Driven Framework for Benchmarking Semantic Web Knowledge Base Systems," by Yuanbo Guo, Abir Qasem, Zhengxiang Pan, and Jeff Heflin, proposes a requirements driven framework for developing benchmarks for Semantic Web knowledge Base Systems (SW KBSs).

c. *Enterprise Information Integration:*

i. "From Wrapping to Knowledge," by José L. Arjona, Rafael Corchuelo, David Ruiz, and Miguel Toro, addresses one of the most challenging problems for Enterprise Information Integration, dealing with heterogeneous information sources on the Web. It proposes an efficient, domain-independent algorithm.

d. *Fluid Web:*

i. "A Component Model and Infrastructure for a Fluid Web," by André Santanchè and Claudia B. Medeiros, presents a solution for the Fluid Web, which allows moving from the document-oriented to a content-oriented perspective, where "content" can be any digital object.

We do believe that this collection of papers is an excellent contribution to the literature of the Semantic Web. We are happy to finalize this special issue and we are really looking forward to your comments. We invite you to work together to develop all of the required bridges between the knowledge and data engineering community and the business world in order to provide the required common ground for exploiting the fascinating technologies of Semantic Web for the promotion of the Knowledge Society. A convergence of computer engineering and management/business strategies will set the Semantics of Business as a key priority for the next years.

ACKNOWLEDGMENTS

Our deepest appreciation and respect goes to Professor Xindong Wu, Editor in Chief of the *IEEE Transactions on Knowledge and Data Engineering*, who gave us the opportunity to serve our community, and for his continuous commitment and contribution to a mutual vision. We wish him health, prosperity, creativity, and well-being. Special thanks to Suzanne Werner, Peer Review Supervisor, for all the great support during the tough development process of this special issue. We would also like to thank the academics and practitioners who contributed their excellent research work to this special issue. Their knowledge, expertise, imagination, and inspiration are evident in every line of this issue. Last, but not least, we are grateful to the 200 reviewers who, with their comments and guidance, helped us to reach an excellent level of quality.

REFERENCES

[1] M. Lytras, "Semantic Web and Information Systems: An Agenda Based on Discourse with Community Leaders," *Int'l J. Semantic Web and Information Systems, Inaugural Issue,* vol. 1, no. 1, pp. i-xii, 2005.

[2] A. Sheth, C. Ramakrishnan, and C. Thomas, "Semantics for the Semantic Web: The Implicit, the Formal, and the Powerful," *Int'l J. Semantic Web and Information Systems,* no. 1, no. 1, pp. 1-18, 2005.

[3] A. Naeve, M. Lytras, W. Nejdl, J. Harding, and N. Balacheff, "Advances of Semantic Web for E-Learning: Expanding Learning Frontiers," *British J. Educational Technology,* vol. 37, no. 3, pp. 321-330, 2006.

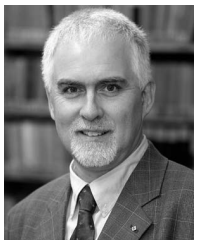
[4] D. Sampson, M. Lytras, G. Wagner, and P. Diaz, "Ontologies and the Semantic Web for E-Learning," *Educational Technology & Soc. J.,* vol. 7, no. 4, pp. 26-28, 2004.

[5] M. Lytras and M. Sicilia, "Knowledge Society a Manifesto for Knowledge and Learning," *Int'l J. Knowledge and Learning,* vol. 1, no. 1, pp. 1-11, 2005.

[6] M. Lytras, M. Sicilia, J. Davies, and V. Kashyap, "Digital Libraries in the Knowledge Era: Knowledge Management and Semantic Web Technologies," *Library Management J.,* vol. 26, nos. 4-5, pp. 170-176, 2005.

[7] A. Naeve, "The Human Semantic Web: Shifting from Knowledge Push to Knowledge Pull," *Int'l J. Semantic Web and Information Systems,* vol. 1, no. 3, pp. 1-30, 2005.

[8] A. Sheth and M. Lytras, *Semantic Web-Based Information Systems: State of the Art Applications (Advances in Semantic Web and Information Systems).* Cybertech Publishing, 2006.



Gottfried Vossen received the master's and PhD degrees as well as the German habilitation in 1981, 1986, and 1990, respectively, all from the Technical University of Aachen in Germany. He is a professor of computer science in the Department of Information Systems and a Director of European Research Center for Information Systems at the University of Muenster in Germany. He has held visiting positions at the University of California in San Diego, at several German universities including the Hasso-Plattner-Institute for Software Systems Engineering in Potsdam near Berlin, at Karlstad University in Sweden, and at The University of Waikato in Hamilton, New Zealand in 2003 and again in 2006. In 2004, he became the European editor-in-chief of Elsevier's *Information Systems—An International Journal* and the director of the European Research Center for Information Systems (ERCIS) in Muenster. His research interests include conceptual as well as application-oriented problems concerning databases, information systems, electronic learning, and the Semantic Web as well as Web 2.0. Dr. Vossen has been a member of numerous program committees of international conferences and workshops. He is an author or coauthor of more than 120 publications and an author, coauthor, or coeditor of 20 books on databases, business process modeling, the Web, e-commerce, and computer architecture.

Miltiadis Lytras received the BSc, MBA, and PhD degrees from the Athens University of Economics and Business. He is an assistant professor in the Department of Computer Engineering and Informatics at the University of Patras, Greece. He serves as the co-editor-in-chief of the *International Journal on Semantic Web and Information Systems* (http://www.idea-group.com/ijswis). He is also the editor-in-chief of the *International Journal of Knowledge and Learning* (http://www.inderscience.com/ijkl) as well as the editor-in-chief of the *International Journal of Teaching and Case Studies* (http://www.inderscience.com/ijtcs). Recently, he undertook the coediting responsibility of the *International Journal of Social and Humanistic Computing* (http://www.inderscience.com/ijshc) as well as the *International Journal of Entertainment Technology and Management* (http://www.inderscience.com/ijenttm). He also serves as the editor-in-chief of *Knowledge and Learning Society Book Series*, http://www.idea-group.-com/requests/details.asp?ID=121, as well as the co-editor-in-chief of the Advances on Semantic Web and Information Systems Book Series, in IDEA Group Publishing. His research focuses on Semantic Web, knowledge management, and e-learning, with more than 80 publications in these areas. He has coedited/coedits 25 special issues in international journals (e.g., *IEEE Internet Computing*, etc.) and has authored/edited 12 books. He is the founder and officer of the Semantic Web and Information Systems Special Interest Group in the Association for Information Systems (http://www.sigsemis.org).

Nick Koudas received the BTech degree from the University of Patras, Greece, the MSc degree from the University of Maryland at College Park, and the PhD degree from the University of Toronto. He is a faculty member at the Department of Computer Science at the University of Toronto. Before joining the University, he was a senior and principal member of research staff at AT&T Labs Research. He was an associate information editor for *SIGMOD Online* (from 1998 until 2003) and serves as an associate editor for the *Information Systems Journal* (since 2000). He is also an editor for the *IEEE Transactions on Knowledge and Data Engineering*, the *ACM Transactions on the Web*, and the *ACM Journal on Data Quality*. He is also a guest editor for the *IEEE Data Engineering Bulletin*. He received the 1998 IEEE International Conference on Data Engineering Best Paper award. He has served on the program committees of various database conferences, including ACM SIGMOD, VLDB, IEEE ICDE, KDD, and EDBT. Dr. Koudas has taught courses related to data management at the University of Maryland College Park, the University of Toronto, and Columbia University. His research interests include data quality, Web search, text processing, stream data processing, XML databases, relational databases (advanced indexing, query optimization, and advanced query models), and data managtement in networking.