

Preface

The success of the World Wide Web depends on the ability of users to store, process and retrieve digital information regardless of distance boundaries, languages and domains of knowledge. The universality and flexibility of the World Wide Web have also enabled the rapid growth of a variety of new services and applications based on human-machine interaction. The semantics of exchanged information and services should be useful not only for human to human communications, but also in that machines would be able to understand and automatically process web content. Semantics give well-defined meaning to web content and enable computers and people to work in cooperation. Today, the crucial challenge becomes the development of languages to express information in a machine processable format. Now more than ever, new advanced techniques and intelligent approaches are required to transform the Web into a universal reasoning and computing machine. Web intelligence attempts to deal with this challenge by exploiting information technologies and artificial intelligence approaches to design the next generation of web-empowered systems and services.

1 Introduction

The semantic Web plays a crucial role in the development of information technologies and services on the World Wide Web. It takes on new challenges in which the meaning of information enable computers to understand the Web content and imitate human intelligence in performing more of the tedious tasks involved in finding, sharing, and combining information on the web. Until now computers have not been able to fully accomplish these tasks without human intervention since web pages are designed to be understood by people, not machines. Tim Berners-Lee originally expressed his vision about the semantic Web. In his word, he wrote: “[I] have a dream for the Web [in which computers] become capable of analyzing all the data on the Web—the content, links, and transactions between people and computers. A ‘Semantic Web’, which should make this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by

machines talking to machines. The ‘intelligent agents’ people have touted for ages will finally materialize.” Significant research efforts attempt to support this vision and propose a set of design principles, formal specifications, and languages specifically designed for a huge data space. Some of these include Resource Description Framework (RDF), Web Ontology languages (OWL), Semantic-based Query Languages and Logic Models to reason on the structure of knowledge embedded in Web-accessible databases and documents. Despite these initiatives, some of the challenges remain in a bottleneck due to the requirement for automating reasoning systems to deal with inconsistency, vastness, uncertainty, vagueness, and deceit in order to deliver on the promise of the Semantic Web. The discipline of Soft Computing has an evolving collection of methodologies, which aims to exploit tolerance for imprecision, uncertainty, and partial truth to study very complex phenomena: those for which more conventional methods have not yielded low cost and complete solutions. Today, Soft Computing provides an attractive opportunity for developing Web intelligence to represent the ambiguity in human thinking with real life uncertainty, reason on vagueness in ontologies, and makes possible the transition between the Web and its semantic successor. In this context, Semantic Web will enable the emergence of digital ecosystems of software and services delivered by the Internet. It will also extend the Internet with capabilities to reason on its resources and their relationships in order to develop the knowledge-based economy in the 21st century.

Why This Book is Interesting? Industrial and technological demands to develop semantic-driven applications in business, commerce, marketing, finance and education have persuaded academia and scholarly communities across the world to include Web intelligence disciplines in their computer science curriculum. Moreover, many research centers are extensively working on this research field which demonstrates an important interest in building the next Semantic World Wide Web. The number of journals in this area has increased and the number of related conferences organized in the last ten years is overwhelming. However, there are relatively few books about web intelligence and semantic web taking into consideration knowledge discovery, semantic network, ontologies and artificial intelligence techniques such as neural network, fuzzy logic and mining algorithms as a new paradigm. Consequently, the need for a new book meets the increasing demands of academia and research communities and provides advanced techniques and methodologies to help undergraduates, graduates and researchers. The primary target audience for the book includes researchers, scholars, postgraduate students and developers who are interested in exploring various areas and disciplines about how semantic techniques and technologies can bridge the gap between users and applications on the web.

2 Book Organization

This book aims to gather the latest advances and innovative solutions in web intelligence and reporting how future services and web-based applications can gain competitive opportunities by applying different emergent semantic techniques to

real-world problems. The main topics of this edited volume cover various advanced semantic technologies, use tools and software for collaboration and simulations of web intelligence, and finally provide case studies and applications in the field of semantic Web. It presents some of the latest advances of various topics in web intelligence and illustrates how organizations can gain competitive advantages by applying the different emergent techniques in the real-world scenarios. The book contains seventeen self-contained chapters which provide optimal reading flexibility. They are organized into four parts as follows:

- Web, Semantic and Intelligence
- Collaboration, Semantic and Intelligence
- Knowledge, Text, Semantic and Intelligence
- Applications and Case Studies

Part I deals with *Web, Semantic and Intelligence* and consists of four chapters.

In Chap. 1, “*The Dilated Triple*”, Rodriguez et al. present a simple model in which the meaning of a statement is defined in terms of other statements, much like the words of a dictionary are defined in terms of each other. This model serves to strengthen the association between concepts and the contexts in which they are to be understood. It provides a particularly simple means of contextualizing an RDF triple by associating it with related statements in the same graph. This approach, in combination with a notion of graph similarity, is sufficient to select only those statements from an RDF graph, which are subjectively most relevant to the context of the requesting process.

In Chap. 2, “*Semantic Web Technologies and Artificial Neural Networks for Intelligent Web Knowledge Source Discovery*”, Caliusco and Stegmayer present some basic concepts and foundations regarding the new Semantic Web and how it is populated with ontologies and why ontology-matching techniques are needed. The idea of software agents that travel the web carrying query request from users has also been addressed in this Chapter. The web knowledge source discovery task is been explained in detail and some motivating scenarios are introduced. To help users avoid irrelevant web search results and wrong decision making, efficient techniques and approaches for developing web intelligence with capabilities for discovering distributed knowledge source are presented.

Wang et al. in Chap. 3, titled “*Computing Similarity of Semantic Web Services in Semantic Nets with Multiple Concept Relations*” propose a novel approach based on application ontologies to improve the selection of semantic web services. After building application ontology by merging semantic service ontologies, the authors represent this application ontology as a fuzzy-weighted semantic net with multiple ontological concept relations, and calculate formal/compound concept similarity on it. The ontological concept similarity is finally used to calculate similarity of semantic services.

Chapter 4, “*General-Purpose Computing on a Semantic Network Substrate*” by Rodriguez presents a model of general-purpose computing on a semantic network substrate. The concepts presented are applicable to any semantic network representation. In the proposed model, the application programming interface, the run-time

program, and the state of the computing virtual machine are all represented in the Resource Description Framework (RDF). The implementations of the concepts presented provide a computing paradigm that leverages the distributed and standardized representational-layer of the Semantic Web.

Part II consists of four chapters and deals with *Collaboration, Semantic and Intelligence*.

In Chap. 5, “*Agent Technology Meets the Semantic Web: Interoperability and Communication Issues*”, Karanastasi and Matsatsini review the recent research on agent technologies and how this technology can serve the scopes of the Semantic Web project including web Agent and the Characteristics of Multi-Agent Systems, Agent Communication Languages, Knowledge and Query Manipulation Language, FIPA Agent Communication Language, and Ontologies.

In Chap. 6, “*Mining of Semantic Image Content Using Collective Web Intelligence*”, Leung et al. describe an indexing method, whereby the aggregate intelligence of different Web users is continuously transferred to the Web. Such intelligence is codified, reinforced, distilled and shared among users so as to enable the systematic mining and discovery of semantic image contents. The described method allows the collaborative creation of image indexes, which is able to instill and propagate deep knowledge and collective wisdom into the Web concerning the advanced semantic characteristics of Web images.

In Chap. 7, “*Suited Support for Distributed Web Intelligence Cooperative Work*”, Decouchant et al. present the PINAS platform, which provides means for supporting cooperative work on the Web. Using cooperative applications that are built employing the services of this infrastructure, several users can access and modify replicated shared entities in a consistent and controlled way. PINAS provides suited features, such as: user identification, multi-site user definition, user and entity naming, shared entity fragmentation and replication, storage, consistency, and automatic distributed updating. The authors propose seamless extensions to standard Web services that can be fully integrated within the Web environment.

In Chap. 8, “*Web services and Software Agents for Tailorable Groupware Design*”, Cheaib et al. present a new groupware architecture model called UD3 that explicitly introduces the notion of tailor ability in designing collaborative applications. This model is based on the integration of web services and software agent technologies, thus using protocols of each while reinforcing their individual strengths in the context of tailorable groupware design. Web services are dynamically invoked by software agents in order to bring new behaviors, and hence, enhancing the collaboration process by dynamically adapting the services offered in the system to the users’ preferences and not the other way around.

Part III consists of five chapters and focus on *Knowledge, Text, Semantic and Intelligence*.

In Chap. 9, “*Toward Distributed Knowledge Discovery on Grid Systems*”, Khac et al. present a distributed data mining (DDM) system based on Grid environments to execute new distributed data mining techniques on very large and distributed heterogeneous datasets. The architecture and motivation for the design are presented. The authors developed prototypes for each layer of the system to evaluate the system features, test each layer as well as whole framework and building simulation

and DDM test suites. Knowledge map layer, key layer of this system, is integrated in this framework.

In Chap. 10, “*Metamodel of Ontology Learning from Text*”, Wisniewski presents the metamodel of the ontology learning from text. The approach is based on the survey of the existing methods, while evaluation is provided in the form of a reference implementation of the introduced metamodel. The author has applied a qualitative evaluation by implementing some of the current state-of-the-art methods and illustrates how they can be described with a metamodel notation.

Ambiguity is a challenge faced by systems that handle natural language. To assuage the issue of linguistic ambiguities found in text classification, Chap. 11, “*An Analysis of Constructed Categories for Textual Classification Using Fuzzy Similarity and Agglomerative Hierarchical Methods*” by Guelpeli et al. proposes a text categorizer using the methodology of Fuzzy Similarity. The clustering algorithms Stars and Cliques are adopted in the Agglomerative Hierarchical method and they authors identify the groups of texts by specifying some type of relationship rule to create categories based on the similarity analysis of the textual terms.

In Chap. 12, “*Emergent XML Mining: Discovering an Efficient Mapping from XML Instances to Relational Schemas*”, Ishikawa proposes an adaptable approach to discovery of database schemas for well-formed XML data such as EDI, news, and digital libraries, which we interchange, filter, or download for future retrieval and analysis. The generated schemas usually consist of more than one table. Author’s approach controls the number of tables to be divided by use of statistics of XML so that the total cost of processing queries is reduced. To achieve this, three functions namely NULL expectation, Large Leaf Fields, and Large Child Fields are introduced for controlling the number of tables to be divided. The author also describes the concept of short paths contained by generated database schemas and their effects on the performance of query processing.

In Chap. 13, “*XML Based Information Systems and Formal Semantics of Programming Languages*”, Despeyroux illustrates how techniques used to define the formal semantics of programming languages can be used in the context of the Web. The author also explores how techniques used in this context can be used to enforce the quality of information systems.

Part IV consists of four chapters and deals with *Applications and Case Studies*.

In Chap. 14, “*Modeling and Testing of Web Based Systems*”, Cavalli et al. present two methodologies to attain automatic test cases generation: The first uses extended finite state machines to model Web services composition described in BPEL, while the other uses UML to model Web applications. Together with the formal models of the web systems, this chapter presents methods for conformance and non-regression test generation.

As web applications are becoming ever larger, more complex and thus more demanding for their users, there is a growing need for customer support. Very often, it is provided by support centers via phone. However, the media break between browser and phone hampers the common understanding of user and consultant. As a result, support becomes ineffective and expensive, and users get frustrated. Screen sharing solutions are one possible solution for this problem, but they have major

disadvantages like high bandwidth requirements, slow performance and, most importantly, the need for a client-side installation. These drawbacks are addressed by VCS, a concept and system for instant co-browsing, that runs directly within the users browser. It equally allows all participants of a support session to see and navigate the same web page on their screens, being aware of what the other person is currently doing on the page. People can directly interact with each other, jointly complete tasks and solve. The event-based nature of the synchronization approach to be presented further facilitates adaptation, so that users with heterogeneous end devices may collaborate. Niederhausen et al., in Chap. 15, “*Web-Based Support by Thin-Client Co-Browsing*”, present VCS and also discuss the special challenges that this approach entails.

In Chap. 16, “*NetPay Micro-Payment Protocols for Three Networks*”, Xiaoling and Grundy describe the NetPay micro-payment protocol that is actually extended from its original pay-per-click for web content to peer-to-peer networks and mobile device networks. The authors outline the key motivation for NetPay, the basic micro-payment protocol using e-coins and e-wallets, and the three variants of the protocol for different domains.

Chapter 17, “*Enforcing Honesty in Fair Exchange Protocols*” by Alaraj and Munro surveys the field of Fair Exchange Protocols and then presents a special type of protocol between a customer (C) and a merchant (M) that enforces one of them to be honest. It makes minimal use of a Trusted Third Party (TTP). The protocol has the features that it: (1) only comprises three messages to be exchanged between C and M; (2) guarantees strong fairness for both C and M; (3) allows both parties to be sure of the item that they will receive from the other party; and (4) resolves disputes automatically online.

3 Acknowledgments

We would like to thank the authors who provided excellent chapters and timely revisions. We are also grateful for their trust in us and patience during the review process. We would like to express our sincere thanks to the reviewers for their tremendous effort and challenging task of choosing high quality chapters, and their valuable criticism that greatly improved the quality of final chapter versions. The editors also would like to thank, Professor Lakhmi C. Jain, the editor-in-chief of the Advanced Information and Knowledge Processing (AI and KP) series of Springer for editorial assistance and excellent cooperative collaboration to produce this important scientific work. We hope this volume motivates its readers to take the next steps beyond building models to implementing, evaluating, comparing, and extending proposed approaches and applications. We finally hope that readers will share our excitement to present this volume on **Web Intelligence: Advanced Semantic Technologies** and find it useful.

4 About the Editors

Youakim Badr received his Ph.D. in Information Systems from the French National Institute for Applied Sciences in Lyon (INSA of Lyon). In 2004, he joined the faculty of the INSA of Lyon as Assistant Professor of Computer Science. Dr. Badr has worked extensively in the field of coupling XML documents and Object-Relational Databases. Through his research he has acquired skills in fields such as Interoperability, Modeling, System Architectures and Networking, and their application to various domains such as Business Processes, Supply Chains, Productions Systems and Virtual Enterprises. His current academic research interests include systems in both the service sector and ICT. In particular, he studies the ecosystem of services and the multidisciplinary modeling approach to design services through the integration of ICT, strategy and processes. He leads the Service-Oriented Enterprise research team which combines industrial and computer engineering approaches. Dr. Badr is vigorously involved in a series of international conferences. He served as General Co-Chair of ICDIM'07, CSTST'08, Programme Chair of INCOS'09, Track chair of IEEE DEST'10, AINA'10, ICETET'09, ICITST'08 and International Program Member of IAS'08, SITIS'07, JFO'07, WCNC'07 and ECWS'06. He is a professional member of ACM, IEEE Services Computing Community, MIL-RLabs/France coordinator, IEEE-SMC Technical Committee on Soft Computing, Digital Ecosystems Community, a member of OW2 and the Service Sciences working group of the Networked European Software and Services Initiative (NESSI).

Richard Chbeir received his Ph.D. in Computer Science from the INSA of Lyon, France in 2001. He is member of IEEE and ACM since 1999. He is currently an Associate Professor in the Computer Science Department of the Bourgogne University, Dijon, France. His research interests are in the areas of distributed multimedia database management, XML similarity and rewriting, spatio-temporal applications, indexing methods, and multimedia access control models. Dr. Chbeir has published (more than 40 peer-reviewed publications) in international journals and books (IEEE Transactions on SMC, Information Systems, Journal on Data Semantics, Journal of Systems Architecture, etc.), conferences (ER, WISE, SOFSEM, EDBT, ACM SAC, Visual, IEEE CIT, FLAIRS, PDCS, etc.), and has served on the program committees of several international conferences (SOFSEM, AINA, IEEE SITIS, ACM SAC, IEEE ISSPIT, EuroPar, SBBD, etc.). He has been organizing many international conferences and workshops (ACM SAC, ICDIM, CSTST, SITIS, etc.). He is currently the Vice-Chair of the ACM SIGAPP and the Chair of its French Chapter.

Ajith Abraham received his Ph.D. degree in Computer Science from Monash University, Melbourne, Australia. His research and development experience includes over 18 years in the Industry and Academia. He works in a multidisciplinary environment involving machine intelligence, network security, sensor networks, e-commerce, Web intelligence, Web services, computational grids, data mining, and applications to various real-world problems. He has given more than 30 plenary lectures and conference tutorials in these areas. He authored or coauthored more than 500 publications. He works with the Norwegian University of Science and Technology, Norway and also coordinate the activities of the Machine Intelligence Research

Labs (MIR Labs), which has representation in 47 countries. He is the Co-Chair of the IEEE Systems Man and Cybernetics Society Technical Committee on Soft Computing. He is the founder of several conference series, which are now sponsored by IEEE and also serves the editorial board of over 30 international Journals.

Aboul Ella Hassanien received his B.Sc. with honours in 1986 and M.Sc. degree in 1993, both from Ain Shams University, Faculty of Science, Ain Sham University, Egypt. On September 1998, he received his doctoral degree from the Department of Computer Science, Graduate School of Science and Engineering, Tokyo Institute of Technology, Japan. He is currently a Professor at Cairo University, Faculty of Computer and Information. He has authored/coauthored over 120 research publications in peer-reviewed reputed journals and conference proceedings. He serves on the editorial board and reviewer of number of journals and on the program committee of several international conferences and he has editing/written more than 18 books. He has received the excellence younger researcher award from Kuwait University.

5 List of Reviewers

Ahmed El Oualkadi (Universit Catholique de Louvain, Belgium)
 Akira Asano (Hiroshima University, Japan)
 Alfredo Cuzzocrea (University of Calabria, Italy)
 Andries Engelbrecht (University of Pretoria, South Africa)
 Bernard Grabot (LGP-ENIT, France)
 Carlos Alberto Reyes-Garcia (Instituto Nacional de Astrofisica Optica Y Electronica, Mexico)
 Chi Shen (University of Kentucky, USA)
 Chrisa Tsinaraki (Technical University of Crete, Greece)
 Christine Verdier (University of Grenoble, France)
 Danielle Boulanger (University of Jean Moulin-MODEME, France)
 Deborah Dahl (Conversational Technologies, USA)
 Elizabeth Goldbarg (Federal University of Rio Grande do Norte, Brazil)
 Estevam Hruschka Jr. (Federal University of Sao Carlos, Brazil)
 Etienne Kerre (University of Gent, Belgium)
 Gabriel Luque (University of Malaga, Spain)
 Georgios Ch. Sirakoulis (Democritus University of Thrace, Greece)
 Hiranmay Ghosh (Tata Consultancy Services, India)
 Hiroshi Ishikawa (Shizuoka University, Japan)
 Ignacio Ponzoni (Universidad Nacional del Sur, Argentina)
 James Lu (Emory University, USA)
 Jaroslaw Kozlak (University of Science and Technology Krakow, Poland)
 Javier J. Sanchez-Medina (University of Las Palmas de Gran Canaria, Spain)
 Kazushi Ohya (Tsurumi University, Japan)
 Kubilay Ecerkale (Hava Harp Okulu, Turkey)
 Mario Koeppen (Kyushu Institute of Technology, Japan)
 Mario Ventresca (University of Waterloo, Canada)
 Maytham Safar (Kuwait University, Kuwait)
 Mei-Ling Shyu (University of Miami, USA)

Michael Blumenstein (Griffith University, Gold Coast, Australia)
Monica Chis (University of Cluj-Napoca, Romania)
Nadine Cullot (Bourgogne University, France)
Oscar Corcho (University of Manchester, UK)
Paolo Merialdo (Universita' degli Studi Roma Tre, Italy)
Patrick Siarry (Universit Paris 12, LiSSi, France)
Patrizia Grifoni (National Research Council of Italy, Italy)
Raquel Barco (Universidad de Malaga, Spain)
Sadok Ben Yahia (University of Tunis, Tunisia)
Saravanan Muthaiyah (George Mason University, USA)
Sebastin Lozano (University of Seville, Spain)
Selma Ayse Ozel (Cukurova University, Turkey)
Siti Mariyam Shamsuddin (University Technology of Malaysia, Malaysia)
Thanasis Daradoumis (Open University of Catalonia, Spain)
Thierry Badard (Universite Laval, Canada)
Thomas Hanne (Applied Sciences Northwestern Switzerland University, Switzerland)
Tianrui Li (Southwest Jiaotong University, China)
Tomasz Smolinski (Emory University, USA)
Urszula Markowska-Kaczmar (Wroclaw University of Technology, Poland)
Ying Ding (Indiana University, USA)
Yinghua Ma (Shanghai JiaoTong University, China)
Zhigang Zeng (Wuhan University of Technology, China)
Zhi-Hong Deng (Peking University, China)

Villeurbanne Cedex, France
Dijon Cedex, France
Trondheim, Norway
Safat, Kuwait

Youakim Badr
Richard Chbeir
Ajith Abraham
Aboul-Ella Hassanien

Emergent Web Intelligence: Advanced Semantic
Technologies

Badr, Y.; Chbeir, R.; Abraham, A.; Hassanien, A.-E.
(Eds.)

2010, XVI, 544 p., Hardcover

ISBN: 978-1-84996-076-2