Editorial

This special issue gathers together the revised versions of the best XML papers presented at the Tenth International World Wide Web Conference (WWW10), held in Hong Kong. In a sense, they bear witness to the fact that XML (the eXtensible Markup Language) is nowadays much more than a brilliant promise, but is in fact a mature and successful reality.

The quality of the 13 papers printed here is evident just by reading them, but some figures can better help the reader to understand how strict the selection procedure has been: each of them has been reviewed by at least six different reviewers and the total number of papers submitted to WWW10 was more than 390. Together, these papers represent an exceptional snapshot of the possibilities of XML, and provide a unique high-quality collection of innovative ideas.

The special issue is divided into four parts (I, II, III and IV). After Part I, which is devoted to the theoretical aspects of XML, the next three parts (II, III and IV) show the importance of XML for real-world applications in a variety of useful contexts.

In Part I (XML Theory), XML is analyzed from a theoretical perspective. While XML and its related technologies have well-defined technical specifications that allow them to function and interoperate, their theoretical foundations are less than clear. These papers help to shed some light by providing a formal analysis of some aspects of XML like keys, constraint models, and XML Schema. In "Keys for XML" the authors relate the fundamental concept of a key as found in the database world to keys in the XML world, showing to what extent the parallelism can be brought forth. Keys and relative keys are redefined in the XML context, and their basic properties analyzed in a wide panorama. Continuing in a parallel to the database world, "A Unified Constraint Model for XML" introduces a useful model of integrity constraints for XML (UCM) designed to help formal reasoning on XML structures, and then analyzes in depth its consistency properties and implementation issues. "MSL: A Model for W3C XML Schema" starts to tackle the big problem of creating a formal model for XML Schema. As XML Schema is rather complex; a complete formalization can help a lot to precisely capture every detail without ambiguities.

In Part II (XML Site Management), the power of XML is exploited in the context of site management, where XML is shown to be a very useful tool to successfully handle such tasks as application development, user interaction, and mirroring. The paper "Vinci: A Service-Oriented Architecture for Rapid Development of Web Applications" describes Vinci, a general architecture designed for quick and effective design of Web applications. Its main feature is modularity, which is accomplished by using XML messaging. While Vinci tries to tackle the whole problem of handling the many applications composing a Web site, the next paper, "Building XML Query Forms and Reports with XQForms", focuses instead on the interaction of the site with users. XQForms are next-generation forms, that can be automatically built from separate high-level components, describing how the structure, query and layout parts of a form and its data work. The next paper, "Enabling Full Service Surrogates Using the Portable Channel Representation", tackles instead the problem of site management at-large, when multiple copies of the same sites are needed. To this extent, an XML model of the specification of a server (the Portable Channel Representation, or PCR) is introduced, that allows easy maintenance of service mirrors in a completely portable way.

In Part III (XML Knowledge Representation) XML meets the Semantic Web, Tim Berners-Lee's famous vision of a next-generation Web where knowledge can be expressed on the Web and then automatically processed and reasoned upon. The papers discuss tools able to encode information, ontologies, and distributed annotations, all using W3C's fundamental technology for the Semantic Web, the XML-based Resource Description Framework (RDF). "The Design and Implementation of the Redland RDF Application Framework" introduces Redland, a full-fledged implementation of RDF that allows efficient storage and querying. Redland provides a clean object-oriented interface, that is available for a variety of programming languages. RDF is also the core technology of Annotea, the system described in "Annotea: An Open RDF Infrastructure for Shared Web Annotations". Annotea is a novel framework for Web annotations, and it is fully interoperable, as it is completely built on standard W3C technologies. Next, "Enabling Knowledge Representation on the Web by Extending RDF Schema" tries to go beyond the current RDF architecture, extending RDF Schema with OIL (the Ontology Inference Layer). OIL provides powerful functionalities for the treatment of ontologies, and it has the characteristic feature of extending RDF Schema in a fully backward-compatible way.

Part IV (XML on the Wire) ends this special issue by showing how XML can be successfully employed in all applications where data and information need to be transmitted over the Internet. The various aspects of data extraction, data push, e-commerce, and data compression are covered from an XML perspective, showing the huge potential that this technology can empower. The paper "Effective Web Data Extraction with Standard XML Technologies" describes ANDES, a system that is able to monitor parts of the Web, and extract meaningful data in XML format. Using XML allows focusing on the essential structure of the data, and not getting lost in possible layout variations. The next paper, "Pushing Reactive Services to XML Repositories using Active Rules", investigates the use of XML for push services, where modifications of XML documents can trigger remote actions on the Web:

using XML protocols and so-called active rules, e-services can be developed that reactively respond to the user's actions. While the previous paper considers the broad area of push technology, the paper "An XML Schema Representation for the Communication Design of Electronic Negotiations" deals with the more specific e-commerce setting: electronic negotiations need a common structure and a precise semantics in order to succeed, and the proposed SilkRoad framework addresses these important problems by using XML Schema to provide an interoperable communication channel. Last but not least, the paper "Algorithms and Programming Models for Efficient Representation of XML for Internet Applications" ends the special issue by considering XML from an orthogonal perspective: size. The authors propose various methods that enable successfully compression and decompression of XML data, so that transmission and handling of XML is considerably eased for every application where time and bandwidth is important.

Finally, we would like to heartily thank all the reviewers, whose comments have substantially contributed to this special issue. We hope you enjoy the results of the hard work of the reviewers and the authors.

Massimo Marchiori
Department of Computer Science
University of Venice
Mestre (Venice), VE 30170, Italy
and
The World Wide Web Consortium (W3C)
MIT Lab for Computer Science
200 Technology Square
Cambridge, MA 02139, USA

Mary Ellen Zurko
IBM Software Group
Westford, MA 01886, USA

Michael R. Lyu
Computer Science & Engineering Department
The Chinese University of Hong Kong
Shatin, Hong Kong
China

Massimo Marchiori holds a double position as professor of Computer Science at the University of Venice, and research architect at MIT in the World Wide Web Consortium (W3C). He received an M.S. in Mathematics and a Ph.D in Computer Science. He has been recipient of several awards for innovative research in the fields of modularity, rewriting system, neural networks, search engines, semantic web, complex systems. He has been working on XML and innovative Web technologies for several years. He has been an appreciated invited speaker at many conferences and symposia, and been many times Chair or Program Committee member. Among various other responsibilities, he is the editor of the world standard for privacy on the Web (P3P), and W3C's responsible for the XML-Query project, which is developing the world standard for querying XML.

Mary Ellen Zurko is the security architect for Lotus brand Next Generation Web Services in the IBM Software Group. She began working on Web security in 1994, when she and another researcher at The Open Group Research Institute integrated DCE security into the Web. She defined the field of "user-centered security" and co-authored an award winning paper on high assurance secure operating systems. She was Program Chair of the 10th International World Wide Web Conference, has been involved in the series since WWW2, and is a member of the organizational committee for that series (IW3C2).

Michael R. Lyu is Professor at the Computer Science and Engineering Department of the Chinese University of Hong Kong. Before that, he worked at the Jet Propulsion Laboratory, at the University of Iowa, at Bell Communications Research (Bellcore), and at Bell Labs (AT&T, and later Lucent Technologies). Dr. Lyu's research interests include software reliability engineering, distributed systems, fault-tolerant computing, wireless communication networks, Web technologies, digital library, and e-commerce systems. He has published over 100 journal and conference papers, participated in more than 30 industrial projects, and helped to develop many commercial systems and software tools. He has been frequently invited as a keynote or tutorial speaker to conferences and workshops all over the world, and he has been Chair and Program Committee member of many conferences. He is an associated editor of IEEE Transactions on Reliability, IEEE Transactions on Knowledge and Data Engineering, and Journal of Information Science and Engineering. Dr. Lyu received a B.S. in Electrical Engineering from National Taiwan University, an M.S. in Computer Engineering from University of California, Santa Barbara, and a Ph.D. in Computer Science from University of California, Los Angeles.