Data-Centric Systems and Applications

Vipul Kashyap · Christoph Bussler ·
Matthew Moran

The Semantic Web

Semantics for Data and Services on the Web

With 61 Figures and 18 Tables

Springer

Vipul Kashyap
Partners HealthCare System
Clinical Informatics R&D
93 Worcester St, Suite 201
Wellesley, MA 02481
USA
vkashyap1@partners.org

Christoph Bussler
Merced Systems, Inc.
333 Twin Dolphin Drive
Suite 500
Redwood Shores, CA 94065
USA
chbussler@aol.com

Matthew Moran
Nortel
Mervue Industry Park
Galway
Ireland
maitiu_moran@hotmail.com

ISBN 978-3-540-76451-9 e-ISBN 978-3-540-76452-6

DOI 10.1007/978-3-540-76452-6

Library of Congress Control Number: 2008931474

ACM Computing Classification (1998): H.4, I.2, J.3

Cover design: KünkelLopka GmbH, Heidelberg

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Preface

A decade ago Tim Berners-Lee proposed an extraordinary vision: despite the phenomenal success of the Web, it would not, and could not, reach its full potential unless it became a place where automated processes could participate as well as people. This meant the publication of documents and data to the web in such a way that they could be interpreted, integrated, aggregated and queried to reveal new connections and answer questions, rather than just browsed and searched. Many scoffed at this idea, interpreting the early emphasis on language design and reasoning as AI in new clothes. This missed the point. The Grand Challenge of the Semantic Web is one that needs not only the information structure of ontologies, metadata, and data, but also the computational infrastructure of Web Services, P2P and Grid distributed computing and workflows. Consequently, it is a truly whole-system and multi-disciplinary effort.

This is also an initiative that has to be put into practice. That means a pragmatic approach to standards, tools, mechanisms and methodologies, and real, challenging examples. It would seem self-evident that the Semantic Web should be able to make a major contribution to clinical information discovery. Scientific communities are ideal incubators: knowledge-driven, fragmented, diverse, a range of structured and unstructured resources with many disconnected suppliers and consumers of knowledge. Moreover, the clinicians and biosciences have embraced the notions of annotation and classification using ontologies for centuries, and have demanding requirements for trust, security, fidelity and expressivity.

This book is the first to describe comprehensively the two main characteristics of the Semantic Web – its information and its processes – and to apply it not to toy, artificial examples but to a challenging application that matters, namely translational medicine. As such, it will become a key text for all of those serious about discovering the many facets of the Semantic Web, those who need to understand the current state of the art as it really is in practice, and those who need to be knowledgeable about its future.

Professor Carole Goble

University of Manchester

Acknowledgments

Vipul would like to dedicate this book to his son Ashish and daughter Aastha. A significant portion of this book was written during paternity leave, when Ashish was born! Vipul also expresses his gratitude to his wife Nitu for tolerating the demands on their family time. The support of Vipul's parents, Yogendra Singh and Vibha, during these times was very valuable and helped reduce stress for Vipul and his family. Vipul would like to acknowledge Dr. Tonya Hongsermeier for her spirited evangelism of the Semantic Web at Partners Healthcare and the W3C Healthcare and Life Sciences Interest Group (HCLSIG). She helped create opportunities to work on Semantic-Web-related projects and pointed me to the translational medicine use case, discussed in Chapter 2 of the book. Vipul also valued his interactions with Dr. Howard Goldberg, which helped him understand pragmatic applications of Semantic Web technologies in the healthcare area. Dr. Blackford Middletonís support for work related to Semantic Web technologies at Partners is also gratefully acknowledged. Discussions and debates with Olivier Bodenreider at the National Library of Medicine were very instructive and introduced Vipul to the area of biomedical informatics. The interactions with various colleagues in HCLSIG, such as Eric Neumann, Kei Cheung, Bill Bug, Susie Stephens, Don Doherty, June Kinoshita, Alan Ruttenberg and others, were also very thought provoking and intellectually stimulating. With Stefan Decker, Vipul had the pleasure of co-presenting a tutorial on the Semantic Web at VLDB 2003, on which this book is based. Vipul has also enjoyed tremendously the follow-on collaboration with Christoph Bussler and Matt Moran on writing this book. Finally, Vipul would like to acknowledge Dr. Amit Sheth, who has been a great mentor and was responsible for introducing me to the research area of semantics and the Semantic Web.

Chris wants to acknowledge a few people he met during his Semantic Web Services journey in Ireland. First there is Dieter Fensel, who in a nutshell was instrumental in opening up the field of semantics to him during many years of collaboration. Then there is the WSMX team that started making Semantic Web Services a reality in various research projects and through a still ongoing open source implementation that started to make its way into standards and industry: David Aiken, Emilia Cimpian, Enrica Dente, Doug Foxvog, Armin Haller, Thomas Haselwanter, Juan Miguel Gomez, Mick Kerrigan, Edward Kilgariff, Adrian Mocan, Matt Moran, Manuel Ohlendor, Eyal Oren, Ina O'Murchu, Brahmananda Sapkota, Laurentiu Vasiliu, Jana Viskova, Tomas Vitvar, Maciej Zaremba, and Michal Zaremba. In context of this book, Vipul and Matt were great co-authors and

Chris enjoyed working together with them a lot during this project. Last but not least he would like to give a huge thanks to his family, Barbara and Vernon, for joining him and supporting him during their time in Ireland; the memories about that time shall remain with them forever.

Matt wants to thank the folks who have encouraged him in his transition into research over the last four years. The WSMX team that Chris has already listed has been fantastic to work with. Matt enjoyed being in the team and thinks its energy and positive approach has been a strong factor in keeping up the momentum in Semantic Web Services research. He would like to particularly thank Tomas Vitvar and Maciej Zaremba with whom he has established a great relationship over the last two years. He is very grateful to Chris and Dieter Fensel, who gave him the chance to step into Semantic Web Services research with DERI. He has learned a lot from them both and hopes to keep learning from them in the future. Writing this book with Vipul and Chris has been a great experience for Matt. The back-and-forth conversations have been generous and enlightening. Finally, he would like to thank his sisters Maeve, Claire, Joanne and Mags, and his parents, for their biased praise and keeping him supplied with Irish stew and wild salmon through the course of his working on this book.

The three authors would like to thank the team at Springer Verlag for its support, especially Ralf Gerstner, who was always engaged, supportive, patient as well as available for discussion and advice.

June 2008

Vipul Kashyap, Boston, MA, USA
Christoph Bussler, San Jose, CA, USA
Matthew Moran, Galway, Ireland

Contents