

Handbook of Semantic Web Technologies

John Domingue, Dieter Fensel, James A. Hendler (Eds.)

Handbook of Semantic Web Technologies

With 203 Figures and 96 Tables

Editors

John Domingue
Knowledge Media Institute
The Open University
Walton Hall
Milton Keynes, MK7 6AA
UK

Dieter Fensel
STI Innsbruck
University of Innsbruck
Technikerstraße 21a
6020 Innsbruck
Austria

James A. Hendler
Department of Computer Science
Rensselaer Polytechnic Institute
Troy, NY 12180
USA

Library of Congress Control Number: 2011921518

ISBN 978-3-540-92912-3

This publication is available also as:

Electronic publication under ISBN 978-3-540-92913-0

Print and electronic bundle under ISBN 978-3-540-92914-7

DOI 10.1007/978-3-540-92913-0

Springer Heidelberg Dordrecht London New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 2011

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Springer is part of Springer Science+Business Media

www.springer.com

Printed on acid-free paper

Foreword

The Semantic Web community has come a long way since its beginnings in the late 1990s and early part of the twenty-first century. There are obviously many ways of categorizing the progress of the topic; however, one easy fit is through three broad phases. Pre Semantic Web, various groups across the globe who had been working in areas related to semantics began to think how semantics and ontologies could aid in certain tasks related to the emerging Web. There then followed significant research funding, which was measured in hundreds of millions of Dollars and Euros, fueling large collaboration projects first in the USA and later in Europe.

Over the last few years, we have seen a second shift. The word “if” disappeared from the vocabulary used in relation to the Semantic Web a long time ago. Now the word “when” has gone too. Primarily around the term Web of Data, we can see commercial take-up by some major players in the area of the Web, Social Networks, and Media. Right now discussions focus more on which particular branch of Semantic Web technology will grow the fastest. We have also seen in Europe a movement away from specific large research calls in the Semantic Web area to semantics being a component within other domains, including, for example, services, media, security, and even networks. This take-up by both research and practitioner communities has been largely based on the fact that the Semantic Web is built upon global standards.

This book began with a realization in the spring 2008 that we seemed to be entering a mainstream phase for our research area, and it thus would be timely to capture the main threads of work. Our first task was of course to determine the book structure.

The book is split into two volumes. The first covers foundational parts and the second applications. The core themes in the first volume comprise semantic annotation and reasoning. By definition, the Semantic Web extends the Web and thus we need semantic extensions of Web languages. This book thus includes chapters on semantic languages for the Web and also mechanisms for inserting semantics into web pages encoded in plain HTML. A main motivation for semantic annotation is that it facilitates machine reasoning. A significant research effort has taken place in relation to this topic, and we cover the primary forms as well as outlining recent efforts to support reasoning at web-scale.

Other foundational issues answer questions such as: how to automatically acquire semantic annotations? how to store these efficiently at large scale? how to query these stores over the Web? and how to express the underlying conceptual structure of the annotations? In the first volume, we also outline the architectural principles underlying the Semantic Web and conclude with some thoughts on its future.

The second volume covers the application of Semantic Web technologies to a number of real-world domains and also to other technical areas. We have seen with the Web how specific application areas can drive innovation, for example, resulting in YouTube and

Facebook for video content and social networking, respectively. The sectors covered in the volume are wide-ranging, incorporating business, science, government, media, broadcasting, and culture. As to the technical areas, we outline here how semantics can improve the management of organizational knowledge and support the use of online services as well as how search is transformed in the context of the Semantic Web.

We would like to take this opportunity to heartily thank all of the chapter authors. This book progressed in a nonlinear fashion, and we are grateful for the patience shown by all and for the rapid response during the most hectic periods. This work was supported by our advisory board and we would like to express our sincere thanks to all of them for their efforts. We are also grateful for the forbearance shown by all our colleagues at Springer.

At the time of writing, we can see that the arrival of the Semantic Web is causing inflection points in a wide variety of ways as it is adopted by new constituents in a number of niches. Real-world requirements and domain-specific opportunities are driving innovation at a growing rate. We believe that this book will serve as a useful reference point for the principles and technologies underlying the Semantic Web as it continues to enter the mainstream, and we eagerly await the results of this process.

John Domingue, Dieter Fensel, and James A. Hendler
May 2011

Advisory Board

V. Richard Benjamins

Telefonica R & D
Madrid
Spain

John Davies

Future Business Applications
and Services
BT Innovate and Design
British Telecommunications Plc
Ipswich
UK

Tim Finin

Department of Computer Science and
Electrical Engineering
University of Maryland
Baltimore, MD
USA

Fausto Giunchiglia

Department of Information and
Communication Technology
University of Trento
Povo
Italy

Mark Greaves

Vulcan Inc.
Seattle, WA
USA

Frank van Harmelen

Department of Artificial Intelligence
VU University Amsterdam
Amsterdam
The Netherlands

Ian Horrocks

Oxford University Computing Laboratory
Oxford
UK

Riichiro Mizoguchi

The Institute of Scientific and Industrial
Research
Osaka University
Ibaraki, Osaka
Japan

Mark A. Musen

The Stanford Biomedical Informatics
Research
Stanford University
Stanford, CA
USA

Guus Schreiber

Department of Computer Science
Vrije Universiteit Amsterdam
Amsterdam
The Netherlands

Daniel Schwabe

Department of Informatics
Pontifícia Universidade Católica do Rio
de Janeiro
Rio de Janeiro
Brasil

Amit P. Sheth

Department of Computer Science &
Engineering
Wright State University
Dayton, OH
USA

Chris Welty

IBM Watson Research Center
Yorktown Heights, NY
USA

Rudi Studer

FZI Research Center for Information
Technology
D-76131 Karlsruhe
Germany
and
Institut für Angewandte Informatik und
Formale Beschreibungsverfahren (AIFB)
Karlsruhe Institute of Technology (KIT)
D-76128 Karlsruhe
Germany

Table of Contents

List of Contributors **xiii**

Volume 1

Part 1 Foundations and Technologies

1 Introduction to the Semantic Web Technologies	3
<i>John Domingue · Dieter Fensel · James A. Hendler</i>	
2 Semantic Web Architecture	43
<i>Andreas Harth · Maciej Janik · Steffen Staab</i>	
3 Semantic Annotations and Retrieval: Manual, Semiautomatic, and Automatic Generation	77
<i>Kalina Bontcheva · Hamish Cunningham</i>	
4 Semantic Annotation and Retrieval: RDF	117
<i>Fabien L. Gandon · Reto Krummenacher · Sung-Kook Han · Ioan Toma</i>	
5 Semantic Annotation and Retrieval: Web of Hypertext – RDFa and Microformats	157
<i>Ben Adida · Mark Birbeck · Ivan Herman</i>	
6 Semantic Annotation and Retrieval: Web of Data	191
<i>Tom Heath · Christian Bizer</i>	
7 Storing the Semantic Web: Repositories	231
<i>Atanas Kiryakov · Mariana Damova</i>	
8 Querying the Semantic Web: SPARQL	299
<i>Emanuele Della Valle · Stefano Ceri</i>	
9 KR and Reasoning on the Semantic Web: OWL	365
<i>Ian Horrocks · Peter F. Patel-Schneider</i>	

10	KR and Reasoning on the Semantic Web: RIF	399
	<i>Michael Kifer</i>	
11	KR and Reasoning on the Semantic Web: Web-scale Reasoning	441
	<i>Spyros Kotoulas · Frank van Harmelen · Jesse Weaver</i>	
12	Social Semantic Web	467
	<i>John G. Breslin · Alexandre Passant · Denny Vrandečić</i>	
13	Ontologies and the Semantic Web	507
	<i>Stephan Grimm · Andreas Abecker · Johanna Völker · Rudi Studer</i>	
14	Future Trends	581
	<i>Lyndon Nixon · Raphael Volz · Fabio Ciravegna · Rudi Studer</i>	

Volume 2

Part 2 Semantic Web Applications

15	Semantic Technology Adoption: A Business Perspective	621
	<i>V. Richard Benjamins · Mark Radoff · Mike Davis · Mark Greaves · Rose Lockwood · Jesús Contreras</i>	
16	Semantic Web Search Engines	659
	<i>Mathieu d'Aquin · Li Ding · Enrico Motta</i>	
17	eScience	701
	<i>Jun Zhao · Oscar Corcho · Paolo Missier · Khalid Belhajjame · David Newmann · David de Roure · Carole A. Goble</i>	
18	Knowledge Management in Large Organizations	737
	<i>John Davies · Paul Warren</i>	
19	eBusiness	787
	<i>Christoph Grün · Christian Huemer · Philipp Liegl · Dieter Mayrhofer · Thomas Motal · Rainer Schuster · Hannes Werthner · Marco Zapletal</i>	
20	eGovernment	849
	<i>Nigel Shadbolt · Kieron O'Hara · Manuel Salvadores · Harith Alani</i>	
21	Multimedia, Broadcasting and eCulture	911
	<i>Lyndon Nixon · Stamatia Dasiopoulou · Jean-Pierre Evain · Eero Hyvönen · Ioannis Kompatsiaris · Raphael Troncy</i>	

22 Semantic Web Services 977
Carlos Pedrinaci · John Domingue · Amit P. Sheth

Glossary1039
Index1053



List of Contributors

Andreas Abecker

FZI Research Center
for Information Technology
Haid-und-Neu-Str. 10-14
D-76131 Karlsruhe
Germany
and
disy Informationssysteme GmbH
Erbprinzenstr. 4-12
D-76133 Karlsruhe
Germany
abecker@fzi.de
Andreas.Abecker@fzi.de

Ben Adida

Creative Commons
San Francisco
USA
and
Center for Research
on Computation and Society
Harvard University
Massachusetts Hall
Cambridge, MA 02138
USA
ben@adida.net

Harith Alani

Knowledge Media Institute
The Open University
Walton Hall
Milton Keynes, MK7 6AA
UK
h.alani@open.ac.uk

Khalid Belhajjame

School of Computer Science
University of Manchester
Oxford Road
Manchester, M13 9PL
UK
Khalid.Belhajjame@manchester.ac.uk

V. Richard Benjamins

Telefonica R&D
Ronda de la Comunicación s/n
28050 Madrid
Spain
rbenjamins@tid.es

Mark Birbeck

WebBackplane
2nd Floor, 69/85 Tabernacle Street
London, EC2A 4RR
UK
mark.birbeck@webbackplane.com

Christian Bizer

Web-based Systems Group
Freie Universität Berlin
Garystr. 21
D-14195 Berlin
Germany
chris@bizer.de

Kalina Bontcheva

Department of Computer Science
University of Sheffield
Regent Court
211 Portobello Street
Sheffield, S1 4DP
UK
K.Bontcheva@dcs.shef.ac.uk

John G. Breslin

Digital Enterprise Research Institute (DERI)
National University of Ireland
Lower Dangan
Galway
Ireland
and
Electrical and Electronic Engineering
School of Engineering and Informatics
National University of Ireland
University Road
Galway
Ireland
john.breslin@nuigalway.ie

Stefano Ceri

Dipartimento
di Elettronica e Informazione
Politecnico di Milano
Via Ponzio 34/5
20133 Milano
Italy
stefano.ceri@polimi.it

Fabio Ciravegna

Department of Computer Science
University of Sheffield
Regent Court
211 Portabello Street
Sheffield, S1 4DP
UK
f.ciravegna@dc.shef.ac.uk

Jesús Contreras

iSOCO
Av. del Partenón
16-18, 1° 7ª Campo de las Naciones
28042 Madrid
Spain
jcontreras@isoco.com

Oscar Corcho

Ontology Engineering Group
Departamento de Inteligencia Artificial
Facultad de Informática
Universidad Politécnica de Madrid
Campus de Montegancedo s/n
Boadilla del Monte,
28660 Madrid
Spain
ocorcho@fi.upm.es

Hamish Cunningham

Department of Computer Science
University of Sheffield
Regent Court
211 Portobello Street
Sheffield, S1 4DP
UK
hamish@wrongonomy.net
hamish@dc.shef.ac.uk

Mariana Damova

Ontotext AD
Sirma Group Corp
135 Tsarigradsko Chaussee
Sofia 1784
Bulgaria
mariana.damova@ontotext.com

Stamatia Dasiopoulou

Informatics and Telematics Institute (ITI)
Centre for Research and Technology
Hellas (CERTH)
6Km Charilaou-Thermi Road
P.O. Box 60361
57001 Thermi-Thessaloniki
Greece
dasiop@iti.gr

John Davies

Future Business Applications
and Services
BT Innovate and Design
British Telecommunications Plc
Ipswich
UK
john.nj.davies@bt.com

Mike Davis

Ovum
119 Farringdon Road
London, EC1R 3ER
UK
mike.davis@ovum.com

Mathieu d'Aquin

Knowledge Media Institute
The Open University
Walton Hall
Milton Keynes, MK7 6AA
UK
m.daquin@open.ac.uk

David de Roure

Oxford e-Research Centre
University of Oxford
7 Keble Road
Oxford, OX1 3QG
UK
dder@ecs.soton.ac.uk

Emanuele Della Valle

Dipartimento
di Elettronica e Informazione
Politecnico di Milano
Via Ponzio 34/5
20133 Milano
Italy
emaunele.dellavalle@polimi.it

Li Ding

Tetherless World Constellation,
Computer Science Department
Rensselaer Polytechnic Institute
110 8th St.
Troy, NY 12180
USA
dingl@cs.rpi.edu

John Domingue

Knowledge Media Institute
The Open University
Walton Hall
Milton Keynes, MK7 6AA
UK
j.b.domingue@open.ac.uk

Jean-Pierre Evain

European Broadcasting Union (EBU)
L'Ancienne-Route 17A
CH-1218 Grand-Saconnex
Switzerland
evain@ebu.ch

Dieter Fensel

STI Innsbruck
University of Innsbruck
Technikerstraße 21a
6020 Innsbruck
Austria
dieter.fensel@sti2.at

Fabien L. Gandon

INRIA-Edelweiss
2004 rt des Lucioles BP93
06902 Sophia Antipolis
France
Fabien.Gandon@sophia.inria.fr

Carole A. Goble

School of Computer Science
University of Manchester
Oxford Road
Manchester, M13 9PL
UK
carole.goble@manchester.ac.uk

Mark Greaves

Vulcan Inc.
505 Fifth Ave South, Suite 900
Seattle, WA 98104
USA
markg@vulcan.com

Stephan Grimm

FZI Research Center
for Information Technology
Haid-und-Neu-Str. 10-14
D-76131 Karlsruhe
Germany
grimm@fzi.de

Christoph Grün

Electronic Commerce Group
Institute for Software Technology
and Interactive System
Vienna University of Technology
Favoritenstrasse 9-11/188
1040 Vienna
Austria
christoph.gruen@ec.tuwien.ac.at

Sung-Kook Han

Semantic Technology Institute
Universität Innsbruck
6020 Innsbruck
Austria
sung-kook.han@sti2.at

Andreas Harth

AIFB Karlsruhe Institute of Technology
(KIT)
Englerstr. 11
D-76131 Karlsruhe
Germany
harth@kit.edu

Tom Heath

Talis Systems Ltd
Knights Court
Solihull Parkway
Birmingham Business Park
Birmingham, B37 7YB
UK
tom.heath@talis.com

James A. Hendler

Department of Computer Science
Rensselaer Polytechnic Institute
Troy, NY 12180
USA
hendler@cs.rpi.edu

Ivan Herman

World Wide Web Consortium and CWI
P.O. Box 94079
Kruislaan 413
1090 GB Amsterdam
The Netherlands
ivan@w3.org

Ian Horrocks

Oxford University Computing Laboratory
Wolfson Building
Parks Road
Oxford, OX1 3QD
UK
Ian.Horrocks@comlab.ox.ac.uk

Christian Huemer

Business Informatics Group
Vienna University of Technology
Karlsplatz 13
1040 Vienna
Austria
christian.huemer@tuwien.ac.at

Eero Hyvönen

Semantic Computing Research Group
(SeCo)
Aalto University
and
University of Helsinki
P.O. Box 15500
FI-00076 Aalto
Finland
evain@ebu.ch

Maciej Janik

Institute for Web Science and
Technologies - WeST
University of Koblenz-Landau
Universitätsstraße 1
56070 Koblenz
Germany
janik@uni-koblenz.de

Michael Kifer

Department of Computer Science
State University of New York
at Stony Brook
Stony Brook, NY 11794-4400
USA
kifer@cs.stonybrook.edu

Atanas Kiryakov

Ontotext AD
Sirma Group Corp
135 Tsarigradsko Chaussee
Sofia 1784
Bulgaria
naso@ontotext.com

Ioannis Kompatsiaris

Informatics and Telematics Institute (ITI)
Centre for Research and Technology
Hellas (CERTH)
6Km Charilaou-Thermi Road
P.O. Box 60361
57001 Thermi-Thessaloniki
Greece
ikom@iti.gr

Spyros Kotoulas

Artificial Intelligence Department
VU University Amsterdam
De Boelelaan 1081a
1081 HV Amsterdam
The Netherlands
kot@few.vu.nl

Reto Krummenacher

Semantic Technology Institute
Universität Innsbruck
6020 Innsbruck
Austria
reto.krummenacher@sti2.at

Philipp Liegl

Business Informatics Group
Vienna University of Technology
Karlsplatz 13
1040 Vienna
Austria
philipp.liegl@tuwien.ac.at

Rose Lockwood

Technology Industry Analyst
24 Harrington Square
London, NW1 2JJ
UK
semanticrose@hotmail.com

Dieter Mayrhofer

Business Informatics Group
Vienna University of Technology
Karlsplatz 13
1040 Vienna
Austria
dieter.mayrhofer@tuwien.ac.at

Paolo Missier

School of Computer Science
University of Manchester
Oxford Road
Manchester, M13 9PL
UK
missier@cs.man.ac.uk

Thomas Motal

Electronic Commerce Group
Institute for Software Technology
and Interactive System
Vienna University of Technology
Favoritenstrasse 9-11/188
1040 Vienna
Austria
thomas.motal@tuwien.ac.at

Enrico Motta

Knowledge Media Institute
The Open University
Walton Hall
Milton Keynes, MK7 6AA
UK
e.motta@open.ac.uk

David Newmann

School of Electronics
and Computer Science
University of Southampton
Southampton, SO17 1BJ
UK
drn@ecs.soton.ac.uk

Lyndon Nixon

Semantic Technology Institute (STI)
International
Amerlingstrasse 19/35
1060 Vienna
Austria
lyndon.nixon@sti2.org

Kieron O'Hara

Intelligence, Agents, Multimedia Group
School of Electronics
and Computer Science
University of Southampton, Highfield
Southampton, SO17 1BJ
UK
kmo@ecs.soton.ac.uk

Alexandre Passant

Digital Enterprise Research Institute
(DERI)
National University of Ireland
Lower Dangan
Galway
Ireland
alexandre.passant@deri.org

Peter F. Patel-Schneider

Alcatel-Lucent Bell Labs Research
Murray Hill
New Jersey
USA
pfps@research.bell-labs.com

Carlos Pedrinaci

Knowledge Media Institute
The Open University
Milton Keynes, MK7 6AA
UK
c.pedrinaci@open.ac.uk

Mark Radoff

Ovum
119 Farringdon Road
London, EC1R 3ER
UK
mark.radoff@ovum.com

Amit P. Sheth

Department of Computer Science
& Engineering
Wright State University
3640 Colonel Glenn Hwy.
Dayton, OH 45435-0001
USA
amit.sheth@wright.edu

Manuel Salvadorés

Intelligence, Agents, Multimedia Group
School of Electronics
and Computer Science
University of Southampton
Highfield
Southampton, SO17 1BJ
UK
ms8@ecs.soton.ac.uk

Steffen Staab

Institute for Web Science
and Technologies - WeST
University of Koblenz-Landau
Universitätsstraße 1
56070 Koblenz
Germany
staab@uni-koblenz.de

Rainer Schuster

Electronic Commerce Group
Institute for Software Technology
and Interactive System
Vienna University of Technology
Favoritenstrasse 9-11/188
1040 Vienna
Austria
Rainer.Schuster@tuwien.ac.at

Rudi Studer

FZI Research Center for Information
Technology
D-76131 Karlsruhe
Germany
and
Institut für Angewandte Informatik und
Formale Beschreibungsverfahren (AIFB)
Karlsruhe Institute of Technology (KIT)
D-76128 Karlsruhe
Germany
rudi.studer@kit.edu
studer@aifb.uni-karlsruhe.de

Nigel Shadbolt

Intelligence, Agents, Multimedia Group
School of Electronics
and Computer Science
University of Southampton, Highfield
Southampton, SO17 1BJ
UK
nrs@ecs.soton.ac.uk

Ioan Toma

Semantic Technology Institute
Universität Innsbruck
6020 Innsbruck
Austria
ioan.toma@sti2.at

Raphaël Troncy

EURECOM
2229 Route des Crêtes
06560 Sophia Antipolis
France
raphael.troncy@eurecom.fr

Frank van Harmelen

Department of Artificial Intelligence
VU University Amsterdam
De Boelelaan 1081a
1081 HV Amsterdam
The Netherlands
Frank.van.Harmelen@cs.vu.nl

Johanna Völker

KR & KM Research Group
University of Mannheim
B6, 26
D-68159 Mannheim
Germany
voelker@informatik.uni-mannheim.de

Raphael Volz

Volz Innovation GmbH
Postfach 45
76597 Loffenau
Germany
rv@volzinnovation.com

Denny Vrandečić

Institut für Angewandte Informatik und
Formale Beschreibungsverfahren – AIFB
Universität Karlsruhe
D-76128 Karlsruhe
Germany
denny.vrandecic@kit.edu

Paul Warren

Programme Manager
Eurescom GmbH
Wieblinger Weg 19/4
D-69123 Heidelberg
Germany
warren@eurescom.eu

Jesse Weaver

Tetherless World Constellation
Rensselaer Polytechnic Institute
Troy
USA
weavej3@rpi.edu

Hannes Werthner

Electronic Commerce Group
Institute for Software Technology
and Interactive System
Vienna University of Technology
Favoritenstrasse 9-11/188
1040 Vienna
Austria
hannes.werthner@tuwien.ac.at

Marco Zapletal

Electronic Commerce Group
Institute for Software Technology
and Interactive System
Vienna University of Technology
Favoritenstrasse 9-11/188
1040 Vienna
Austria
marco.zapletal@tuwien.ac.at

Jun Zhao

Department of Zoology
Oxford University
South Parks Road
Oxford, OX1 3PS
UK
jun.zhao@zoo.ox.ac.uk