The Epistemology of Intelligent Semantic Web Systems

Synthesis Lectures on the Semantic Web: Theory and Technology

Editor
Ying Ding, ***Indiana University***
Paul Groth, ***Elsevier Labs***

Synthesis Lectures on the Semantic Web: Theory and Application is edited by Ying Ding of Indiana University and Paul Groth of Elsevier Labs. Whether you call it the Semantic Web, Linked Data, or Web 3.0, a new generation of Web technologies is offering major advances in the evolution of the World Wide Web. As the first generation of this technology transitions out of the laboratory, new research is exploring how the growing Web of Data will change our world. While topics such as ontology-building and logics remain vital, new areas such as the use of semantics in Web search, the linking and use of open data on the Web, and future applications that will be supported by these technologies are becoming important research areas in their own right. Whether they be scientists, engineers or practitioners, Web users increasingly need to understand not just the new technologies of the Semantic Web, but to understand the principles by which those technologies work, and the best practices for assembling systems that integrate the different languages, resources, and functionalities that will be important in keeping the Web the rapidly expanding, and constantly changing, information space that has changed our lives.
Topics to be included:

- Semantic Web Principles from linked-data to ontology design
- Key Semantic Web technologies and algorithms
- Semantic Search and language technologies
- The Emerging "Web of Data" and its use in industry, government and university applications
- Trust, Social networking and collaboration technologies for the Semantic Web
- The economics of Semantic Web application adoption and use
- Publishing and Science on the Semantic Web
- Semantic Web in health care and life sciences

The Epistemology of Intelligent Semantic Web Systems
Mathieu d'Aquin and Enrico Motta
2016

Entity Resolution in the Web of Data
Vassilis Christophides, Vasilis Efthymiou, and Kostas Stefanidis
2015

Library Linked Data in the Cloud: OCLC's Experiments with New Models of Resource Description
Carol Jean Godby, Shenghui Wang, and Jeffrey K. Mixter
2015

Semantic Mining of Social Networks
Jie Tang and Juanzi Li
2015

Social Semantic Web Mining
Tope Omitola, Sebastián A. Ríos, and John G. Breslin
2015

Semantic Breakthrough in Drug Discovery
Bin Chen, Huijun Wang, Ying Ding, and David Wild
2014

Semantics in Mobile Sensing
Zhixian Yan and Dipanjan Chakraborty
2014

Provenance: An Introduction to PROV
Luc Moreau and Paul Groth
2013

Resource-Oriented Architecture Patterns for Webs of Data
Brian Sletten
2013

Aaron Swartz's A Programmable Web: An Unfinished Work
Aaron Swartz
2013

Incentive-Centric Semantic Web Application Engineering
Elena Simperl, Roberta Cuel, and Martin Stein
2013

Publishing and Using Cultural Heritage Linked Data on the Semantic Web
Eero Hyvönen
2012

VIVO: A Semantic Approach to Scholarly Networking and Discovery
Katy Börner, Michael Conlon, Jon Corson-Rikert, and Ying Ding
2012

Linked Data: Evolving the Web into a Global Data Space
Tom Heath and Christian Bizer
2011

© Springer Nature Switzerland AG 2022
Reprint of original edition © Morgan & Claypool 2016

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means—electronic, mechanical, photocopy, recording, or any other except for brief quotations in printed reviews, without the prior permission of the publisher.

The Epistemology of Intelligent Semantic Web Systems

Mathieu d'Aquin and Enrico Motta

ISBN: 978-3-031-79470-4 paperback
ISBN: 978-3-031-79471-1 ebook

DOI 10.1007/978-3-031-79471-1

A Publication in the Morgan & Claypool Publishers series
SYNTHESIS LECTURES ON THE SEMANTIC WEB: THEORY AND TECHNOLOGY

Lecture #14
Series Editors: Ying Ding, *Indiana University*
Paul Groth, *Elsevier Labs*
Founding Editor Emeritus: James Hendler, *Rensselaer Polytechnic Institute*
Series ISSN
Print 2160-4711 Electronic 2160-472X

The Epistemology of Intelligent Semantic Web Systems

Mathieu d'Aquin and Enrico Motta
Knowledge Media Institute, The Open University

SYNTHESIS LECTURES ON THE SEMANTIC WEB: THEORY AND TECHNOLOGY #14

ABSTRACT

The Semantic Web is a young discipline, even if only in comparison to other areas of computer science. Nonetheless, it already exhibits an interesting history and evolution. This book is a reflection on this evolution, aiming to take a snapshot of where we are at this specific point in time, and also showing what might be the focus of future research.

This book provides both a conceptual and practical view of this evolution, especially targeted at readers who are starting research in this area and as support material for their supervisors. From a conceptual point of view, it highlights and discusses key questions that have animated the research community: what does it mean to be a Semantic Web system and how is it different from other types of systems, such as knowledge systems or web-based information systems? From a more practical point of view, the core of the book introduces a simple conceptual framework which characterizes Intelligent Semantic Web Systems. We describe this framework, the components it includes, and give pointers to some of the approaches and technologies that might be used to implement them. We also look in detail at concrete systems falling under the category of Intelligent Semantic Web Systems, according to the proposed framework, allowing us to compare them, analyze their strengths and weaknesses, and identify the key fundamental challenges still open for researchers to tackle.

KEYWORDS

semantic web, linked data, intelligent systems, knowledge engineering, knowledge-based systems, ontologies

Contents

Preface xi

Acknowledgments xiii

1 Characterizing the Semantic Web 1

1.1 It is a Stack After All! 2
1.2 The Web as a Layer on Top of the Internet 6
1.3 Linked Data as a Layer on Top of the Web 7
1.4 The Semantic Web as a Layer on Top of Linked Data 9

2 Anatomy of an Intelligent Semantic Web System 13

2.1 Knowledge-based Systems 13
2.2 Interacting with Knowledge in Knowledge-based Systems 14
2.3 Interacting with the Semantic Web as a Knowledge Base 16
2.4 Intelligent Semantic Web Systems as Views over the Semantic Web 19
2.5 Components of an Intelligent Semantic Web System 21
2.5.1 Discovery 23
2.5.2 Selection 24
2.5.3 Integration 25
2.5.4 Curation 26
2.6 Conclusion 27

3 Exemplar Intelligent Semantics Web Systems 29

3.1 Search and Recommendation 30
3.1.1 Google Knowledge Graph 30
3.1.2 Seevl/dbrec 33
3.1.3 DiscOU 35
3.2 Question Answering 37
3.2.1 IBM Watson 39
3.2.2 PowerAqua 41
3.3 Data Analysis and Sense-Making 43
3.3.1 Garlik Data Patrol 43

3.3.2 Rexplore . . . 46
3.4 Conclusion . . . 47

4 The Challenges that Need to be Addressed to Realize Ubiquitous Intelligent Semantic Web Systems . . . 49
4.1 Technological Challenges: Scale, Robustness, and Distribution . . . 49
4.2 Non-Technological Challenges: The Human in the Loop . . . 53
4.2.1 Quality . . . 53
4.2.2 Policies and Rights . . . 55
4.2.3 Privacy . . . 56
4.2.4 Interaction . . . 58
4.3 Conclusion . . . 60

References . . . 61

Authors' Biographies . . . 73

Preface

In this book, we take both a conceptual and a practical view of the Semantic Web. The Semantic Web is a young discipline, even if compared only to other areas of computer science. Nonetheless, it already exhibits an interesting history and evolution. From a very general vision which, as presented in the foundational article by Berners Lee et al. (2001), integrated different techniques and approaches (agents, databases, knowledge representation, web technologies, etc.), the focus quickly moved to establishing and standardizing a core set of technologies for the representation, distribution, and access to knowledge and data on the Web—e.g., RDF, OWL, SPARQL, etc. Once these core technologies were established, the research community was then able to focus on developing applications and building a large-scale web of data, in accordance with the Linked Data principles.

Having gone in a rather accelerated manner through this cycle, from vision to impact, this is the time to see where this is all going. This is the conceptual aspect of this book: The *Epistemology* of Intelligent Semantic Web Systems. The debate that started as soon as the idea of the Semantic Web was put forward by Tim Berners Lee and colleagues, mostly from the Artificial Intelligence area, has indeed never stopped: What does it mean to be a Semantic Web system? How is it different from other types of systems, such as knowledge systems or web-based information systems? Our goal is not necessarily to provide definitive answers to these questions, but to highlight and discuss in a concise manner the key elements that need to be considered and understood when engaging with this new class of systems.

First, we look at the Semantic Web as part of the Web, and characterize it as a conceptual construct, part of a stack of structures and networks, rather than a stack of technologies (Chapter 1). The idea here is that, by understanding what makes the Semantic Web something of a higher level of abstraction than the Web itself, we have a better view of the implications in terms of the challenges to be addressed, and of the opportunities for new applications and systems to be developed. This view of the Semantic Web naturally leads to thinking of it as a distributed, networked system where knowledge is shared and managed globally. It also leads us to think of it as a type of knowledge system with specific characteristics to do with being distributed, networked, and thus open and without centralized control.

Second, we consider conceptual models established in the field of Knowledge-based Systems and explore how they can be adapted to the open, decentralized nature of the Semantic Web (Chapter 2). In doing so, we establish a simple conceptual framework to characterize Intelligent Semantic Web Systems in terms of the functions and components they include. The goal here is not only to better understand what makes an Intelligent Semantic Web System, but also to establish a reference model useful both to recognize and compare Intelligent Semantic Web

Systems, and also to guide their design. This framework can be seen as defining the core of the book, joining up the conceptual discussion on the nature of the Semantic Web with the more practical view on what it means to develop an Intelligent Semantic Web System. In contrast with other characterizations, here we do not focus on actual technologies but on the various conceptual elements that need to be present when designing Intelligent Semantic Web Systems.

The practical aspect of this book is therefore in this focus on Intelligent Semantic Web *Systems*, and in the deliberately rapid jump from debating the nature of the Semantic Web, to looking at what it means concretely. Hence, once we have established the general, conceptual framework, we look in detail at concrete systems that, according to this framework, fall into the category of Intelligent Semantic Web Systems (Chapter 3). The goal here is that, by looking at concrete examples, we can see how the general model captured by the Intelligent Semantic Web System framework is instantiated in practice. In addition, the model also allows us to compare different systems, thus analyzing their relative strengths and weaknesses and their conformance to our framework.

Having clarified the nature of Intelligent Semantic Web Systems, we are then left with the question: "What next?" Hence, in the final chapter, we look at the key open research challenges that come out of our analysis. To some extent, here we come to two seemingly contradictory conclusions: On the one hand, the area of Intelligent Semantic Web Systems has evolved into a mature field with research directly leading to major opportunities and impacts already clearly visible in many academic and commercial systems. On the other hand, the area is still in its infancy with many fundamental challenges still open for researchers to tackle.

This contradiction is one of the reasons why this book is especially suitable for those (e.g., students) who are starting research in this area and as support material for their supervisors. Much of the material in this book comes from our own experience of more than a decade working as researchers and application developers in the Semantic Web area and from the many discussions on these issues we have had with other members of the research community (see for example d'Aquin et al., 2008a). Much of the thinking here also comes from the experience of establishing in 2003 the first international summer school dedicated to Semantic Web Research (SSSW). SSSW is still today the key educational event in the Semantic Web area and has evolved dramatically since its first edition, reflecting the evolution of the area but also, to some extent, contributing to shaping it. This book is a reflection on this evolution, with the aim to take a snapshot of where we are at this specific point in time, and also to show what the future will be like, or at least should be like!

Mathieu d'Aquin and Enrico Motta
April 2016

Acknowledgments

The content of this book is based on research started by our group more than a decade ago. It derives, directly or indirectly, from the work of the past and present members of this group, namely: Alessandro Adamou, Sofia Angeletou, Elizabeth Cano Basave, Carlo Allocca, Claudio Baldassarre, Emanuele Bastianelli, Enrico Daga, Maurizio di Matteo, Martin Dzbor, Salman Elahi, Miriam Fernandez, Jorge Gracia, Laurian Gridinoc, Davide Guidi, Tom Heath, Yuangui Lei, Ning Li, Shuangyan Liu, Vanessa Lopez, Andriy Nikolov, Francesco Osborne, Michele Pasin, Silvio Peroni, Dnyanesh Rajpathak, Marta Sabou, Angelo Antonio Salatino, Lucia Specia, Keerthi Thomas, Ilaria Tiddi, Victoria Uren, Maria Vargas-Vera, Paul Warren, Fouad Zablith and many others who have been, in one way or another, associated with us. We are immensely grateful to have had the opportunity to work with such brilliant researchers and to debate with them the state, direction, and impact of the semantic web, linked data, and Intelligent Semantic Web Systems. We are also grateful to all our colleagues, who came for a visit or we met at conferences, seminars, and other events, and with whom we discussed and wrote about the challenges and opportunities of the semantic web. Finally, we thank our families and friends for being there and mostly bearing with us.

Mathieu d'Aquin and Enrico Motta
April 2016