

A Developer's Guide to the Semantic Web

Liyang Yu

A Developer's Guide to the Semantic Web

Second Edition



Springer

Liyang Yu
Delta Air Lines
Atlanta
USA

ISBN 978-3-662-43795-7 ISBN 978-3-662-43796-4 (eBook)
DOI 10.1007/978-3-662-43796-4
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014950052

© Springer-Verlag Berlin Heidelberg 2011, 2014, Corrected 3rd printing 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, 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.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To my parents, Zaiyun Du my mother and Hanting Yu my father.

The truth is, they cannot read this dedication without someone translating it. However, this is never a problem for them, since there is something in this world that goes beyond the boundary of all languages and all cultures, and still remains the same to all human hearts. It lets my parents understand every word I have said here without the need of any translation at all.

It is the love they have been giving me. I will never be able to pay them back enough, and I can only wish that I will be their son in my next life, so I can continue to love them, and be loved.

Preface to the First Edition

Objectives of the Book

This book is all about the Semantic Web.

From its basics, the Semantic Web can be viewed as a collection of standards and technologies that allow machines to understand the meaning (semantics) of information on the Web. It represents a new vision about how the Web should be constructed so its information can be processed automatically by machines on a large scale.

This exciting vision opens the possibility of numerous new applications on the Web. Since 2001, there have been many encouraging results in both academic world and real application world. A whole suite of standards, technologies, and related tools have been specified and developed around the concept of the Semantic Web.

However, such an extensive and rapid progress of the Semantic Web has presented a steep learning curve for those who are new to the Semantic Web. Understanding its related concepts, learning the core standards and key technical components, and finally reaching the point where one can put all these into real development work requires a considerable amount of effort.

To facilitate this learning process, a comprehensive and easy-to-follow text is a must. This book, *A Developer's Guide to the Semantic Web*, serves this exact purpose. It provides an in-depth coverage on both the *What-Is* and *How-To* aspects of the Semantic Web. From this book, you will not only obtain a solid understanding about the Semantic Web but also learn how to combine all the pieces together to build new applications on the Semantic Web. More specifically,

- It offers a complete coverage of all the core standards and technical components of the Semantic Web. This coverage includes RDF, RDFS, OWL (both OWL 1 and OWL 2), and SPARQL (including features offered by SPARQL 1.1). Other related technologies are also covered, such as Turtle, Microformats, RDFa, GRDDL, and SKOS.

- It provides an in-depth description of multiple well-known applications and projects in the area of the Semantic Web, such as FOAF, semantic Wiki, SearchMonkey by Yahoo!, Rich Snippets by Google, Open Linked Data Project, and DBpedia Project.
- It explains the key concepts, core standards, and technical components in the context of examples. In addition, the readers will be taken in a step-by-step fashion through the development of each example. Hopefully for the first time ever, such teaching method will ease the learning curve for those who have found the Semantic Web a daunting topic.
- It includes several complete programming projects, which bridge the gap between *What-Is* and *How-To*. These example applications are real coding projects and are developed from the scratch. In addition, the code generated by these projects can be easily reused in the readers' future development work.

Intended Readers

The book is written with the following readers in mind:

- Software engineers and developers who are interested in learning the Semantic Web technology in general.
- Web application developers who have the desire and/or needs to study the Semantic Web and build semantic Web applications.
- Researchers working in research institutes who are interested in the Semantic Web research and development.
- Undergraduate and graduate students from computer science departments, whose focus of work is in the area of the Semantic Web.
- Practitioners in related engineering fields. For example, Data Mining engineers whose work involves organizing and processing a large amount of data by machines.

The prerequisites needed to understand this book includes the following:

- Working knowledge of Java programming language.
- Basic understanding of the Web, including its main technical components such as URL, HTML, and XML.

Structure of the Book

This book is organized as follows:

- Chapters 1–6 covers the basic concept, the core standards, and technical components of the Semantic Web. The goal of these chapters is to show you the *What-Is* aspect about the Semantic Web.

Chapter 1 introduces the concept of the Semantic Web by using a simple example. With this example, the difference between the traditional Web and the Semantic Web is clearly revealed. Further discussion in this chapter helps you to establish a solid understanding about the concept of the Semantic Web.

Chapter 2 covers RDF in great detail to give you a sound technical foundation to further understand the Semantic Web. If you are not familiar with RDF, you should not skip this chapter, since everything else is built upon RDF. In addition, Turtle format is presented in this chapter, which will be used to understand the material presented in Chap. 6.

Chapter 3 goes on with other RDF-related technologies, including Microformats, RDFa and GRDDL. If you prefer to get a full picture about the Semantic Web as quickly as possible, you can skip this chapter. However, the material presented in this chapter will be necessary in order to understand Chap. 8.

Chapter 4 presents RDF Schema and also introduces the concept of ontology. You should not skip this chapter since Chap. 5 is built upon this chapter. SKOS is also presented in this chapter; you can skip it if you are not working with any existing vocabularies in knowledge management field.

Chapter 5 discusses OWL in great detail and covers both OWL 1 and OWL 2. This is one of the key chapters in this book and should not be skipped. Unless RDF Schema can satisfy the needs of your application, you should spend enough time to understand OWL, which will give you the most updated information about latest ontology development language.

Chapter 6 covers SPARQL. This is another chapter that you should carefully read. Working on the Semantic Web without using SPARQL is like working with database systems without knowing SQL. Notice SPARQL 1.1 is covered in this chapter as well. At the time of this writing, SPARQL 1.1 has not become a standard yet, so when you are reading this book, notice the possible updates.

- Chapters 7–11 provides an in-depth discussion of some well-known semantic Web applications/projects in the real application world. These chapters serve as a transition from knowing *What-Is* to understanding *How-To* in the world of the Semantic Web.

Chapter 7 presents FOAF (Friend Of A Friend) project. The FOAF ontology is arguably the most widely used ontology at this point. The goal of this chapter is to introduce you to a real world example in the social networking area. Since the modeling of this domain does not require any specific domain knowledge, it is easy to follow and you can therefore focus on appreciating the power of the Semantic Web. This chapter should not be skipped, not only because of the popularity of the FOAF ontology but also because this ontology has been used frequently in the later chapters as well.

Chapter 8 presents Google's Rich Snippets and Yahoo!'s searchMonkey; both are using RDFa and Microformats as the main tools when adding semantic markups. These are important examples, not only because they are currently the

major semantic Web applications developed by leading players in the field, but also they show us the benefits of having the added semantics on the Web.

Chapter 9 discusses the topic of Semantic Wiki, together with a real world example. This chapter represents the type of the Semantic Web applications built by using manual semantic markup. After reading this chapter, you should not only see the power of the added semantics but also start to understand those situations where manual semantic markup can be a successful solution.

Chapter 10 presents DBpedia in great detail. DBpedia is a well-known project in the Semantic Web community, and a large number of real world semantic Web applications take advantage of the DBpedia datasets directly or indirectly. Also, DBpedia gives an example of automatic semantic markup. Together with Chap. 9, where manual semantic markup is used, you have a chance to see both methods at work.

Chapter 11 discusses the Linked Open Data project (LOD) as an real world implementation example of the Web of Data concept. For the past several years, LOD has attracted tremendous attention from both the academic world and real application world. In fact, DBpedia, as a huge dataset, stays in the center of LOD cloud. Therefore, LOD together with DBpedia, becomes a must for anyone who wants to do development work on the Semantic Web. More specifically, this chapter covers both the production and the consumption aspects of Linked Data; it also provides application examples that are built upon LOD cloud. In addition, this chapter explains how to access LOD programmatically, which should be very useful to your daily development work.

- Chapters 12–15 is the section of *How-To*. After building a solid foundation for development work on the Semantic Web, this section presents three different running applications that are created from scratch. The methods, algorithms, and concrete classes presented in these chapters will be of immediate use to your future development work.

Chapter 12 helps to build a foundation for your future development work on the Semantic Web. More specifically, it covers four major tool categories you should know, namely, development frameworks, reasoners, ontology engineering tools, and other tools such as search engines for the Semantic Web. This chapter also discusses some related development methodology for the Semantic Web, such as the Ontology-driven Software Development Methodology. Furthermore, since ontology development is the key of this methodology, this chapter also presents an ontology development guide that you can use.

Chapter 13 covers a popular development framework named Jena, to prepare you for your future development work on the Semantic Web. More specifically, this chapter starts from how to setup Jena development environment and then presents a Hello World example to get you started. In what follows, this chapter covers the basic operation every semantic Web application needs, such as creating RDF models, handling persistence, querying RDF dataset, and

inferencing with ontology models. After reading this chapter, you will be well prepared for real development work.

Developing applications for the Semantic Web requires a set of complex skills, and this skill set lands itself on some basic techniques. In Chap. 13, you have learned some basics. Chapter 14 continues along the same path by building an agent that implements the Follow-Your-Nose algorithm on the Semantic Web. After all, most semantic Web applications will have to be based on the Web, so moving or crawling from one dataset to another on the Web with some specific goals in mind is a routine task. Follow-Your-Nose method is one such basic technique. Besides implementing this algorithm, Chap. 14 also introduces some useful operations, such as how to remotely access SPARQL endpoints.

Chapter 15 presents two additional semantic Web applications from scratch. The first application helps you to create an e-mail list that you can use to enhance the security of your e-mail system. The second one is a ShopBot that runs on the Semantic Web, and you can use it to find products that satisfy your own specific needs. These two projects are both discussed in great detail, showing how applications on the Semantic Web are built. This includes RDF documents handling, ontology handling, inferencing based on ontologies, and SPARQL query handling, just to name a few.

Where to Get the Code

The source code for all the examples and application projects in this book can be downloaded from the author's personal Web site, www.liyangyu.com.

Acknowledgment

My deepest gratitude goes to a remarkable person in my life, Dr. Weirong Ding, for supporting me in all the ways that one can ever wish to be supported. It is not nearly as possible to list all the supports she gave me, but her unreserved confidence in my knowledge and talents has always been a great encouragement for me to finish this book. Being the first reader of this book, she has always been extremely patient with many of my ideas and thoughts, and interestingly enough, her patience has made her a medical doctor who is also an expert of the Semantic Web. And to make the readers of this book become experts of the Semantic Web, I would like to share something she always says to me: "never give yourself excuses and always give 200% of yourself to reach what you love."

I would like to thank Dr. Jian Jiang, a good friend of mine, for introducing me to the field of the Semantic Web, for many interesting and insightful discussions along the road of this book.

My gratitude is also due to Mr. Ralf Gerstner, Senior Editor at Springer. As a successful IT veteran himself, he has given me many valuable suggestions about the content and final organization of this book. The communication with him is always quite enjoyable: it is not only prompt and efficient but also very insightful and helpful. It is simply my very good luck to have a chance to work with an editor like Ralf.

Finally, I would like to express my love and gratitude to my beloved parents, for their understanding and endless love. They always give me the freedom I need, and they accept my decisions even when they cannot fully understand them. In addition, I wanted to thank them for being able to successfully teach me how to think and speak clearly and logically when I was at a very young age, so I can have one more dream fulfilled today.

Delta Air Lines, Atlanta, USA

Liyang Yu

Preface to the Second Edition

In the three years since the first edition of this book was published, I have received numerous email messages from readers all over the world commenting on the book and suggesting how it could be improved. Meanwhile, the Semantic Web world has been developing and experiencing exciting changes and improvements, noticeably some new standards such as SPARQL 1.1 and RDB2RDF, and also new developments such as schema.org and important real-world applications that built upon schema.org. In addition, I have also built up a large file of ideas based on my own research in the area and my own experience gained when building semantic Web applications for a variety of organizations. It is indeed a good time to update the book and make all these new material available to our readers.

The most obvious changes in this second edition are the new chapters described as follows.

- *schema.org and Semantic Markup (Chap. 10)*. schema.org has been extremely important and popular since the day it was launched (June of 2011). Developed by Google, Yahoo, and Bing, it is considered as the first mainstream support of the vision of the Semantic Web. This new chapter covers everything about schema.org, including its background and rationale, its vocabulary, and the markup languages recommended. It also includes two examples to showcase some real-world applications built upon schema.org (Google-rich snippets and LRMI Project).
- *Social Network and the Semantic Web (Chap. 11)*. It is not surprising that most of us are more or less related to the Web by participating in some kind of social network sites. It might however be surprising to realize that the Semantic Web technology has actually been playing a key role in these social network sites. This chapter focuses on this topic and helps to understand how the Semantic Web technology has been changing the way social networking sites work. More specifically, this chapter uses three most popular social networking sites, namely, Facebook, Twitter, and Pinterest as examples, and examines their *semantic components* in great detail. For Facebook, we take a look at the *Open Graph protocol*; for Twitter, we study *Twitter cards*; and for Pinterest, we focus

on *rich pins*. This not only shows how the idea of the Semantic Web can help social networking sites, but also serves as examples to the developers, so they can draw inspiration and come up with their own applications.

- *Other Recent Applications: data.gov and Wikidata (Chap. 12)*. This chapter uses two more recent developments to further illustrate how the idea of the Semantic Web can be applied to the Web and data that surround us. The first example is data.gov. We discuss the background of data.gov, how it is related to the Semantic Web, and examples are included to show the benefits of using the Semantic Web technologies on government open data. The second example is wikidata, a popular project that has been constantly under the spotlight recently. The relationship between Wikipedia, DBpedia, and wikidata is first discussed, followed by a close look at the semantic components inside wikidata. This will be another eye-opening project to the readers, because the Semantic Web components used in wikidata have indeed changed how Wikipedia is constructed and maintained.
- *Getting Started: Change Your Data Into Structured Data (Chap. 13)*. This chapter is motivated by questions from our readers. For example, if one is not consuming public RDF data, how should one create his/her own RDF content? If all existing structured data are stored in database tables, how should this structured information be converted into RDF content? Which ontology should be used? Is there a way to understand a complex ontology that is created by someone else? This chapter attempts to answer all these questions, in preparation for the readers to start their own development work. A main focus of this chapter is the RDB2RDF W3C standard, which is discussed in great detail and example implementation is also represented in a way that the readers can directly follow it in their own development work.
- *A Search Engine that Supports Rich Snippets (Chap. 17)*. This chapter is added as another example of developing semantic Web applications. It is important because (1) it directly shows how the Semantic Web idea can be used to enhance the performance of a search engine and (2) the implementation in this chapter can be directly adapted to build customized search engines that support rich snippets for different organizations.

Among the new material in existing chapters, Chap. 6, *SPARQL: Querying the Semantic Web*, has been greatly enhanced by covering the language features of the new standard, SPARQL 1.1, more thoroughly and completely. Compared to the first edition of this book, the coverage of SPARQL 1.1 in this second edition has changed from 14 to more than 40 pages.

Besides the above enhancement, most existing chapters are updated with new links, new figures if necessary, and new version numbers if applicable.

Finally, this second edition is organized into three parts. Part I, *Core of the Semantic Web*, containing Chaps. 1–6, covers the foundation of the Semantic Web, Part II, *Applied Semantic Web*, containing Chaps. 7–12, describes some application examples and latest development in the area of the Semantic Web, and Part III, *Building Your Own Applications on the Semantic Web*, containing Chaps. 13–18,

offers concrete development guidelines and detailed descriptions of the necessary technical foundations, together with real projects and coding examples. This new layout clearly shows how the whole book is organized, with the goal of helping the readers to more easily conquer the learning curve needed to master the world of the Semantic Web.

Delta Air Lines, Atlanta, USA

Liyang Yu

Contents

Part I Core of the Semantic Web

1	A Web of Data: Toward the Idea of the Semantic Web	3
1.1	A Motivating Example: Data Integration on the Web	4
1.1.1	A Smart Data Integration Agent	4
1.1.2	Is Smart Data Integration Agent Possible?	10
1.1.3	The Idea of the Semantic Web	11
1.2	A More General Goal: A Web Understandable to Machines . . .	12
1.2.1	How Do We Use the Web?	12
1.2.2	What Stops Us From Doing More?	15
1.2.3	Again, the Idea of the Semantic Web	17
1.3	The Semantic Web: A First Look	17
1.3.1	The Concept of the Semantic Web	17
1.3.2	The Semantic Web, Linked Data and the Web of Data . . .	18
1.3.3	Some Basic Things About the Semantic Web	20
	Reference	21
2	The Building Block for the Semantic Web: RDF	23
2.1	RDF Overview	23
2.1.1	RDF In Official Language	23
2.1.2	RDF in Plain English	25
2.2	The Abstract Model of RDF	29
2.2.1	The Big Picture	29
2.2.2	Statement	30
2.2.3	Resource and Its URI Name	31
2.2.4	Predicate and Its URI Name	36
2.2.5	RDF Triples: Knowledge That Machines Can Use	38
2.2.6	RDF Literals and Blank Node	40
2.2.7	A Summary So Far	47

2.3	RDF Serialization: RDF/XML Syntax	48
2.3.1	The Big Picture: RDF Vocabulary	48
2.3.2	Basic Syntax and Examples	49
2.3.3	Other RDF Capabilities and Examples	66
2.4	Other RDF Sterilization Formats	73
2.4.1	Notation-3, Turtle and N-Triples	73
2.4.2	Turtle Language	73
2.5	Fundamental Rules of RDF	81
2.5.1	Information that Is Understandable by Machines	81
2.5.2	Distributed Information Aggregation	84
2.5.3	A Hypothetical Real World Example	85
2.6	More About RDF	88
2.6.1	Dublin Core: Example of Predefined RDF Vocabulary	88
2.6.2	XML vs. RDF?	90
2.6.3	Use a RDF Validator	93
2.7	Summary	94
3	Other RDF-Related Technologies: Microformats, RDFa and GRDDL	97
3.1	Introduction: Why Do We Need These?	97
3.2	Microformats	98
3.2.1	Microformats: The Big Picture	98
3.2.2	Microformats: Syntax and Examples	99
3.2.3	Microformats and RDF	105
3.3	RDFa	106
3.3.1	RDFa: The Big Picture	106
3.3.2	RDFa Attributes and RDFa Elements	107
3.3.3	RDFa: Rules and Examples	108
3.3.4	RDFa and RDF	115
3.4	GRDDL	116
3.4.1	GRDDL: The Big Picture	116
3.4.2	Using GRDDL with Microformats	117
3.4.3	Using GRDDL with RDFa	118
3.5	Summary	119
4	RDFS and Ontology	121
4.1	RDFS Overview	121
4.1.1	RDFS in Plain English	121
4.1.2	RDFS in Official Language	123
4.2	RDFS + RDF: One More Step Toward Machine-Readable	123
4.2.1	A Common Language to Share	123
4.2.2	Machine Inferencing Based on RDFS	125
4.3	RDFS Core Elements	126
4.3.1	The Big Picture: RDFS Vocabulary	126
4.3.2	Basic Syntax and Examples	127
4.3.3	Summary So Far	146

4.4	The Concept of Ontology	150
4.4.1	What Is Ontology	151
4.4.2	The Benefits of Ontology	151
4.5	Building the Bridge to Ontology: SKOS	152
4.5.1	Knowledge Organization Systems (KOS)	152
4.5.2	Thesauri vs. Ontologies	154
4.5.3	Filling the Gap: SKOS	156
4.6	Another Look at Inferencing Based on RDF Schema	163
4.6.1	RDFS Ontology Based Reasoning: Simple, Yet Powerful	163
4.6.2	Good, Better and Best: More Is Needed	166
4.7	Summary	166
5	OWL: Web Ontology Language	169
5.1	OWL Overview	169
5.1.1	OWL in Plain English	169
5.1.2	OWL in Official Language: OWL 1 and OWL 2	170
5.1.3	From OWL 1 to OWL 2	172
5.2	OWL 1 and OWL 2: The Big Picture	173
5.2.1	Basic Notions: Axiom, Entity, Expression and IRI Names	173
5.2.2	Basic Syntax Forms: Functional-Style, RDF/XML Syntax, Manchester Syntax and XML Syntax	174
5.3	OWL 1 Web Ontology Language	175
5.3.1	Defining Classes: The Basics	176
5.3.2	Defining Classes: Localizing Global Properties	178
5.3.3	Defining Classes: Using Set Operators	188
5.3.4	Defining Classes: Using Enumeration, Equivalent and Disjoint	191
5.3.5	Our Camera Ontology So Far	194
5.3.6	Define Properties: The Basics	197
5.3.7	Defining Properties: Property Characteristics	203
5.3.8	Camera Ontology Written Using OWL 1	212
5.4	OWL 2 Web Ontology Language	216
5.4.1	What Is New in OWL 2	217
5.4.2	New Constructs for Common Patterns	217
5.4.3	Improved Expressiveness for Properties	221
5.4.4	Extended Support for Datatypes	232
5.4.5	Punning and Annotations	237
5.4.6	Other OWL 2 Features	241
5.4.7	OWL Constructs in Instance Documents	246
5.4.8	OWL 2 Profiles	250
5.4.9	Our Camera Ontology in OWL 2	256
5.5	Summary	262

6	SPARQL: Querying the Semantic Web	265
6.1	SPARQL Overview	265
6.1.1	SPARQL in Official Language	265
6.1.2	SPARQL in Plain Language	266
6.1.3	RDF Datasets and SPARQL Endpoints	267
6.2	SPARQL 1.0 Query Language	269
6.2.1	The Big Picture	271
6.2.2	SELECT Query	274
6.2.3	CONSTRUCT Query	302
6.2.4	DESCRIBE Query	305
6.2.5	ASK Query	306
6.2.6	What Is Missing from SPARQL 1.0?	307
6.3	SPARQL 1.1 Query Language	308
6.3.1	Introduction: What Is New?	308
6.3.2	SPARQL 1.1 Query	309
6.3.3	SPARQL 1.1 Federated Query	327
6.3.4	SPARQL 1.1 Update	330
6.3.5	Other SPARQL 1.1 Features	342
6.4	Summary	352

Part II Applied Semantic Web

7	FOAF: Friend of a Friend	357
7.1	What FOAF Is and What It Does	357
7.1.1	FOAF in Plain English	357
7.1.2	FOAF in Official Language	358
7.2	Core FOAF Vocabulary and Examples	359
7.2.1	The Big Picture: FOAF Vocabulary	360
7.2.2	Core Terms and Examples	361
7.3	Create Your FOAF Document and Get into the Friend Circle	368
7.3.1	How Does the Circle Work?	369
7.3.2	Create Your FOAF Document	371
7.3.3	Get into the Circle: Publish Your FOAF Document	371
7.3.4	From Web Pages for Human Eyes to Web Pages for Machines	374
7.4	Semantic Markup: A Connection Between the Two Worlds	375
7.4.1	What Is Semantic Markup?	376
7.4.2	Semantic Markup: Procedure and Example	376
7.4.3	Semantic Markup: Feasibility and Different Approaches	380
7.5	Summary	382

8	DBpedia	383
8.1	Introduction to DBpedia	383
8.1.1	From Manual Markup to Automatic Generation of Annotation	383
8.1.2	From Wikipedia to DBpedia	384
8.1.3	The Look-and-Feel of DBpedia: Page Redirect	385
8.2	Semantics in DBpedia	389
8.2.1	Infobox Template	389
8.2.2	Creating DBpedia Ontology	392
8.2.3	Infobox Extraction Methods	398
8.3	Accessing DBpedia Dataset	401
8.3.1	Using SPARQL to Query DBpedia	401
8.3.2	Direct Download of DBpedia Datasets	406
8.3.3	Access DBpedia as Linked Data	412
8.4	Summary	414
	Reference	414
9	Linked Open Data	415
9.1	The Concept of Linked Data and Its Basic Rules	415
9.1.1	The Concept of Linked Data	415
9.1.2	How Big Are the Web of Linked Data and the LOD Project?	417
9.1.3	The Basic Rules of Linked Data	418
9.2	Publishing RDF Data on the Web	419
9.2.1	Identifying Things with URIs	420
9.2.2	Choosing Vocabularies for RDF Data	431
9.2.3	Creating Links to Other RDF Data	433
9.2.4	Serving Information as Linked Data	440
9.3	The Consumption of Linked Data	446
9.3.1	Discover Specific Targets on the Linked Data Web	448
9.3.2	Accessing the Web of Linked Data	452
9.4	Linked Data Application	462
9.4.1	Linked Data Application Example: Revyu	462
9.4.2	Web 2.0 Mashups vs. Linked Data Mashups	470
9.5	Summary	472
10	schema.org and Semantic Markup	475
10.1	Introduction to schema.org	475
10.1.1	What Is schema.org?	475
10.1.2	Understanding the schema.org Vocabulary	477
10.2	Content Markup Using schema.org	479
10.2.1	RDFa 1.1 Lite: A Simple Subset of RDFa	479
10.2.2	What Markup Format to Use?	485
10.2.3	Type Checking and Other Issues	486
10.2.4	Validating Your Markup	488

10.3	Content Markup Example 1: Google Rich Snippets	490
10.3.1	What Is Rich Snippets: An Example	490
10.3.2	Google Rich Snippets: Semantic Markup Using schema.org	492
10.3.3	Using Google Rich Snippets Testing Tool	501
10.4	Content Markup Example 2: LRMI Project	505
10.4.1	The Idea of LRMI	505
10.4.2	LRMI Specification	507
10.4.3	LRMI Implementation Examples	510
10.5	Summary	514
	References	515
11	Social Networks and the Semantic Web	517
11.1	Overview of Social Networking Websites	517
11.2	Facebook's Open Graph Protocol	519
11.2.1	Open Graph Protocol	520
11.2.2	How Does It Work: Creating Typed Links Using OGP	524
11.2.3	Implications for the Semantic Web	529
11.3	Twitter Cards for Structured Information	530
11.3.1	Twitter Cards Overview	530
11.3.2	How Does It Work: Structured Information for Rich Tweets	533
11.3.3	Structured Information, But Not Semantic Web Yet	538
11.4	Rich Pins for Structured Information	541
11.4.1	Rich Pin Overview	541
11.4.2	How Does It Work: Generating Rich Pins Using schema.org	543
11.4.3	Semantic Markup at Work	547
11.5	Summary	549
12	Other Recent Applications: data.gov and Wikidata	551
12.1	Data.gov and the Semantic Web	551
12.1.1	Understanding Data.gov	551
12.1.2	How Is Data.gov Related to the Semantic Web?	557
12.1.3	Potential eGov Standards: Breaking the Boundaries of Datasets	561
12.1.4	Example Data.gov Applications	564
12.2	Wikidata and the Semantic Web	566
12.2.1	From Wikipedia to Wikidata	566
12.2.2	Three Phases of the Wikidata Project	571
12.2.3	Wikidata as a Data Repository	574
12.2.4	Wikidata and the Semantic Web	577
12.3	Summary	585

Part III Building Your Own Applications on the Semantic Web

13	Getting Started: Change Your Data into Structured Data	589
13.1	RDF Data in General	589
13.1.1	What Does RDF Data Refer to?	590
13.1.2	Decide in Which Format to Publish Your RDF Data	591
13.1.3	Decide Which Ontology to Use to Publish Your Data	596
13.2	Creating RDF Data Manually	602
13.2.1	Popular Editors and Validators	602
13.2.2	Examples: Using TopBraid to Create RDF Data	603
13.3	RDB2RDF: W3C's Standard for Converting DB Content to RDF Triples	608
13.3.1	RDB2RDF: General Background	608
13.3.2	Direct Mapping from RDB to RDF	609
13.3.3	R2RML: RDB to RDF Mapping You Can Control	613
13.4	RDB2RDF Example Implementation	623
13.4.1	RDB2RDF Direct Mapping	623
13.4.2	Step-by-Step R2RML Example: Virtuoso	624
13.5	Summary	642
14	Building the Foundation for Development on the Semantic Web	643
14.1	Development Tools for the Semantic Web	643
14.1.1	Frameworks for the Semantic Web Applications	643
14.1.2	Reasoners for the Semantic Web Applications	647
14.1.3	Ontology Engineering Environments	650
14.1.4	Other Tools: Search Engines for the Semantic Web	654
14.1.5	Where to Find More?	654
14.2	Semantic Web Application Development Methodology	655
14.2.1	From Domain Models to Ontology-Driven Architecture	655
14.2.2	An Ontology Development Methodology Proposed by Noy and McGuinness	661
14.3	Summary	666
	Reference	667
15	Example: Using Jena for Development on the Semantic Web	669
15.1	Jena: A Semantic Web Framework for Java	669
15.1.1	What Is Jena and What Can It Do for Us?	669
15.1.2	Getting the Jena Package	670
15.1.3	Using Jena in Your Projects	671
15.2	Basic RDF Model Operations	676
15.2.1	Creating an RDF Model	677
15.2.2	Reading an RDF Model	683
15.2.3	Understanding an RDF Model	685

15.3	Handling Persistent RDF Models	692
15.3.1	From In-Memory Model to Persistent Model	692
15.3.2	Setting up MySQL	693
15.3.3	Database-Backed RDF Models	694
15.4	Inferencing Using Jena	702
15.4.1	Jena Inferencing Model	702
15.4.2	Jena Inferencing Examples	703
15.5	Summary	710
16	Follow Your Nose: A Basic Semantic Web Agent	711
16.1	The Principle of Follow-Your-Nose Method	711
16.1.1	What Is the Follow-Your-Nose Method?	711
16.1.2	URI Declarations, Open Linked Data and Follow-Your-Nose Method	713
16.2	A Follow-Your-Nose Agent in Java	714
16.2.1	Building the Agent	714
16.2.2	Running the Agent	721
16.2.3	More Clues for Follow-Your-Nose	724
16.2.4	Can You Follow Your Nose on Traditional Web?	725
16.3	A Better Implementation of Follow-Your-Nose Agent: Using SPARQL Queries	727
16.3.1	In-Memory SPARQL Operation	727
16.3.2	Using SPARQL Endpoints Remotely	732
16.4	Summary	735
17	A Search Engine That Supports Rich Snippets	737
17.1	Why This Is an Interesting Project	737
17.2	Introduction to Lucene	738
17.2.1	Lucene and Our Own Customized Search Engine	738
17.2.2	Core Components of Lucene	739
17.2.3	Use Lucene in Your Development Environment	745
17.3	Preparing the Semantic Markups	746
17.3.1	From Semantic Markup to Rich Snippets	746
17.3.2	Different Deployment Models of the Markup	747
17.3.3	Examples of Markup	749
17.4	Building the Search Engine	752
17.4.1	Creating the Indexer	752
17.4.2	Creating the Searcher	759
17.4.3	Using Web Container to Start the Search	763
17.5	Test It Out and Possible Expansions	768
17.5.1	Test Runs of the Search Engine	768
17.5.2	Possible Expansions	770
17.6	Summary	772

18	More Application Examples on the Semantic Web	773
18.1	Building Your Circle of Trust: A FOAF Agent You Can Use . . .	773
18.1.1	Who Is on Your E-Mail List?	773
18.1.2	The Basic Idea	774
18.1.3	Building the EmailAddressCollector Agent . . .	777
18.1.4	Can You Do the Same for the Traditional Web?	788
18.2	A ShopBot on the Semantic Web	788
18.2.1	A ShopBot We Can Have	788
18.2.2	A ShopBot We Really Want	790
18.2.3	Building Our ShopBot	799
18.2.4	Discussion: From Prototype to Reality	817
18.3	Summary	818
	Index	819