

Programming the Semantic Web

Toby Segaran, Colin Evans, and Jamie Taylor

O'REILLY®

Beijing • Cambridge • Farnham • Köln • Sebastopol • Taipei • Tokyo

Table of Contents

Foreword xi

Preface xiii

Part I. Semantic Data

1. Why Semantics? 3

- Data Integration Across the Web 4
- Traditional Data-Modeling Methods 5
 - Tabular Data 6
 - Relational Data 7
 - Evolving and Refactoring Schemas 9
 - Very Complicated Schemas 11
 - Getting It Right the First Time 12
- Semantic Relationships 14
- Metadata Is Data 16
- Building for the Unexpected 16
- “Perpetual Beta” 17

2. Expressing Meaning 19

- An Example: Movie Data 21
- Building a Simple Triplestore 23
 - Indexes 23
 - The add and remove Methods 24
 - Querying 25
- Merging Graphs 26
- Adding and Querying Movie Data 28
- Other Examples 29
 - Places 29
 - Celebrities 31
 - Business 33

3. Using Semantic Data	37
A Simple Query Language	37
Variable Binding	38
Implementing a Query Language	40
Feed-Forward Inference	43
Inferring New Triples	43
Geocoding	45
Chains of Rules	47
A Word About “Artificial Intelligence”	50
Searching for Connections	50
Six Degrees of Kevin Bacon	51
Shared Keys and Overlapping Graphs	53
Example: Joining the Business and Places Graphs	53
Querying the Joined Graph	54
Basic Graph Visualization	55
Graphviz	55
Displaying Sets of Triples	56
Displaying Query Results	57
Semantic Data Is Flexible	59

Part II. Standards and Sources

4. Just Enough RDF	63
What Is RDF?	63
The RDF Data Model	64
URIs As Strong Keys	64
Resources	65
Blank Nodes	66
Literal Values	68
RDF Serialization Formats	68
A Graph of Friends	69
N-Triples	70
N3	72
RDF/XML	73
RDFa	76
Introducing RDFLib	80
Persistence with RDFLib	83
SPARQL	84
SELECT Query Form	86
OPTIONAL and FILTER Constraints	87
Multiple Graph Patterns	89
CONSTRUCT Query Form	91

ASK and DESCRIBE Query Forms	91
SPARQL Queries in RDFLib	92
Useful Query Modifiers	94
5. Sources of Semantic Data	97
Friend of a Friend (FOAF)	97
Graph Analysis of a Social Network	101
Linked Data	105
The Cloud of Data	106
Are You Your FOAF file?	107
Consuming Linked Data	110
Freebase	116
An Identity Database	117
RDF Interface	118
Freebase Schema	119
MQL Interface	121
Using the metaweb.py Library	123
Interacting with Humans	125
6. What Do You Mean, "Ontology"?	127
What Is It Good For?	127
A Contract for Meaning	128
Models Are Data	128
An Introduction to Data Modeling	129
Classes and Properties	129
Modeling Films	132
Reifying Relationships	134
Just Enough OWL	135
Using Protégé	140
Creating a New Ontology	140
Editing an Ontology	141
Just a Bit More OWL	145
Functional and Inverse Functional Properties	146
Inverse Properties	146
Disjoint Classes	146
Keepin' It Real	148
Some Other Ontologies	148
Describing FOAF	148
A Beer Ontology	149
This Is Not My Beautiful Relational Schema!	152
7. Publishing Semantic Data	155
Embedding Semantics	155

Microformats	156
RDFa	158
Yahoo! SearchMonkey	160
Google's Rich Snippets	161
Dealing with Legacy Data	162
Internet Video Archive	162
Tables and Spreadsheets	167
Legacy Relational Data	169
RDFLib to Linked Data	172

Part III. Putting It into Practice

8. Overview of Toolkits	183
Sesame	183
Using the Sesame Java API	184
RDFS Inferencing in Sesame	193
A Servlet Container for the Sesame Server	196
Installing the Sesame Web Application	196
The Workbench	197
Adding Data	199
SPARQL Queries	200
REST API	202
Other RDF Stores	203
Jena (Open Source)	204
Redland (Open Source)	204
Mulgara (Open Source)	204
OpenLink Virtuoso (Commercial and Open Source)	204
Franz AllegroGraph (Commercial)	205
Oracle (Commercial)	205
SIMILE/Exhibit	205
A Simple Exhibit Page	206
Searching, Filtering, and Prettier Views	209
Linking Up to Sesame	211
Timelines	212
9. Introspecting Objects from Data	215
RDFObject Examples	215
RDFObject Framework	217
How RDFObject Works	225
10. Tying It All Together	227
A Job Listing Application	227

Application Requirements	228
Job Listing Data	228
Converting to RDF	228
Loading the Data into Sesame	231
Serving the Website	232
CherryPy	232
Mako Page Templates	233
A Generic Viewer	234
Getting Data from Sesame	236
The Generic Template	236
Getting Company Data	237
Crunchbase	238
Yahoo! Finance	241
Reconciling Freebase Connections	243
Specialized Views	244
Publishing for Others	248
RDFa	248
RDF/XML	250
Expanding the Data	251
Locations	251
Geography, Economy, Demography	252
Sophisticated Queries	253
Visualizing the Job Data	255
Further Expansion	258

Part IV. Epilogue

11. The Giant Global Graph	261
Vision, Hype, and Reality	262
Participating in the Global Graph Community	264
Releasing Data into the Commons	265
License Considerations	266
The Data Cycle	267
Bracing for Continuous Change	268
Index	271