

Explorer's Guide to the Semantic Web

THOMAS B. PASSIN



MANNING

Greenwich
(74° w. long.)

contents

preface xiii
acknowledgments xv
about this book xvii

1

The Semantic Web 1

1.1 What is the Semantic Web? 3

Indexing and retrieving information 5 ▪ *Meta data* 5
Annotation 6 ▪ *A huge interoperable database* 6
Machine retrieval of data 6 ▪ *Services* 7 ▪ *Discovery* 7
Intelligent agents 8

1.2 Two Semantic Web scenarios 8

Can the Semantic Web work this way? 10

1.3 The Semantic Web's foundation 10

Resources 11 ▪ *Standardized addressing* 11 ▪ *Small set of commands* 12 ▪ *Scalability and large networks* 12
Openness, completeness, and consistency 13 ▪ *The Web and the Semantic Web* 13

1.4 The Semantic Web layer cake 14

The base 16 ▪ *Properties and relationships* 17 ▪ *Analysis, verification, and trust* 17

1.5 Summary 18

2

Describing data with RDF 19**2.1 Introducing RDF 22**

Some terminology 22 ■ Identifying resources 23 ■ Anonymous resources 25 ■ RDF and conventional databases 25

2.2 Working with RDF properties 28

*Properties as resources 28 ■ Names, labels, and identifiers 29
Properties of statements 30*

2.3 Visualizing statements with RDF graphs 30

*Resources with many statements 31 ■ Anonymous nodes 32
Resources as objects of statements 33 ■ Container nodes 34
Graphing properties of statements 35*

2.4 Six RDF applications 37

*RDF in Mozilla 37 ■ RSS 37 ■ RDF for annotations:
Annotea 38 ■ Bibliographic meta data: Dublin Core 39
WebScriber: fusing information 39 ■ Haystack: personal
information management 40*

2.5 Meshing data and meta data 41

*The data model 41 ■ The data in table format 42
The relationships between tables 43 ■ The RDF version of
the data 44 ■ Table versions of the data 46 ■ Why bother
with RDF? 47*

2.6 Sharing data 47

RDF/XML syntax 48 ■ Non-XML formats 56

2.7 RDF in the real world of the Web 56

*What does a URI reference indicate? 56 ■ Contradictory
statements 58 ■ Incomplete information 58
Limitations 58*

2.8 Summary 59

3

Navigating information with topic maps 60**3.1 What are topic maps? 61**

*An example index 62 ■ The subject of a topic 66
Properties 68 ■ Scopes 69 ■ Summary of key features 70
Collocation, navigation, and representation of information 70
Merging topic maps 71 ■ Maturity of Topic Maps software 72*

3.2 Basic models of topic maps 72

Abstract model 73 ■ Implementation approaches 77

- 3.3 Sharing topic maps between computers 78
- 3.4 Topic map examples 79
 - Weather events* 80 ▪ *Conference proceedings* 81
- 3.5 Topic maps and the Web 82
- 3.6 Topic maps and RDF 84
 - RDF: information by the atom* 84 ▪ *Topic maps: proteins of knowledge* 85 ▪ *The subject, revisited* 85 ▪ *Theoretical basis* 86 ▪ *Data structures as first-class citizens* 86
 - Strengths and weaknesses* 86
- 3.7 How topic maps are used in practice 87
- 3.8 Summary 88

4

Annotation 90

- 4.1 What is annotation? 91
- 4.2 Annotations at full power 92
- 4.3 Current web annotation systems 94
 - Wiki collaboratives* 95 ▪ *Annotea* 96 ▪ *Multivalent browser* 99
- 4.4 Improving annotation 103
- 4.5 Summary 105

5

Searching 106

- 5.1 Searching the Web 108
 - Kinds of searches* 108 ▪ *So near and yet so far* 109
 - Behind the scenes* 110
- 5.2 Search strategies 112
 - Keywords* 112 ▪ *Ontologies* 113 ▪ *Meta data* 114
 - Semantic analysis* 114 ▪ *Semantic focusing* 117 ▪ *Social analysis* 118 ▪ *Multiple passes* 120 ▪ *Peer to peer* 120
 - Clustering* 121
- 5.3 Distorting results and spoofing search engines 123
- 5.4 Searching and the Semantic Web 124
 - Self-describing meta data: no panacea* 124 ▪ *Semantic Web possibilities for improving searching* 124 ▪ *Searching and web services* 125
- 5.5 Summary 126

6 **The role of logic 127**

6.1 What logic is good for 128

*Rules 129 ■ Inferring facts 130 ■ Explanations 131
Contradictions and interpretations 131 ■ Ontologies 133
The representation of knowledge 134 ■ Queries 134
Combining information 135*

6.2 All logics aren't created equal 135

*First-order logic: the logic of individual things 136
Second-order logic: the logic of types and relationships 136
Finessing complexity 137 ■ Life beyond first-order logic 137*

6.3 Two pitfalls 138

Don't swallow the whole Web 138 ■ Knowledge pollution 139

6.4 Patchwork quilt or seamless whole? 139

6.5 Summary 140

7 **Ontology 141**

7.1 Basic classification 143

Lists, hierarchies, and trees 143 ■ Classification groups 145

7.2 Arranging names 148

Names and identifiers 148 ■ Properties 149

7.3 Building ontologies 150

*Frameworks 151 ■ On designing ontologies 151
Other considerations 152*

7.4 Languages for ontologies 155

RDFS 155 ■ OWL 161 ■ DAML + OIL 168

7.5 Summary 169

8 **Semantic Web services 170**

8.1 What are web services? 171

*Web pages as services 172 ■ Beyond the plain web page 172
How semantic are today's web services? 173 ■ Elements of web
services 173*

8.2 Exchanging data 174

*Web only? 175 ■ To RDF or not? 175 ■ SOAP 175
RDF with SOAP 177 ■ HTML forms 177*

- 8.3 Invoking services 178
 - Using HTTP for web services* 178 ▪ *About HTTP messages* 178 ▪ *Remote procedure calls* 181 ▪ *The Web versus RPC* 181 ▪ *The RPC controversy* 184
- 8.4 Describing and finding services 186
 - Connecting to services* 186 ▪ *Discovering services* 194
 - Describing web service processes* 198
- 8.5 Will web services converge with the Semantic Web? 201
- 8.6 Summary 202

9

Agents 204

- 9.1 What is an intelligent agent? 205
- 9.2 Agents at work 207
- 9.3 Basic agent types 207
 - Logic-based agents* 208 ▪ *Reactive agents* 209
 - Belief-desire-intention agents* 209 ▪ *Layered architectures* 209
- 9.4 Agent interactions 210
- 9.5 Agents on the Semantic Web 212
 - Beyond factory agents* 212 ▪ *Agent evolution* 213
- 9.6 Frameworks and standards 213
 - Foundation for Intelligent Physical Agents* 213 ▪ *FIPA-OS and other FIPA implementations* 215 ▪ *Java Agent Services* 217
- 9.7 Summary 217

10

Distributed trust and belief 219

- 10.1 The territory 221
- 10.2 Tools of trust 222
 - Private keys* 222 ▪ *Public keys* 223 ▪ *Digests* 223
 - Public Key Infrastructure* 223 ▪ *Digital signing* 227
 - A trust scorecard* 228
- 10.3 Reliable and believable? 229
 - All statements aren't equal* 230 ▪ *Handling contradictory information* 230 ▪ *Dealing with information of uncertain reliability* 231

10.4 Agents and the Web of Trust 233

10.5 Summary 234

11

Putting it all together 236

11.1 Just what is the Semantic Web? 237

11.2 Scenarios redux 238

The first scenario 238 ■ The second scenario 242

The scenarios in perspective 244

11.3 Some Key Issues 244

Scalability 245 ■ Ontologies—universal or piecework 245

Identity 246 ■ Strong AI and the role of logical reasoning 247

Embedded semantic markup 247 ■ Web services and the

Semantic Web 248 ■ Trust, belief, and confidence 248

11.4 How semantic will it be? 249

11.5 Can it really come about? 250

*The technology front 250 ■ Plugging in 251 ■ Growing the
Semantic Web 254*

Appendix Case studies 256

A.1 FOAF: Friend of a Friend 257

Sharing a bit of yourself 257 ■ A FOAF example 258

*Identification without universal IDs 259 ■ Board of a
board 260 ■ Lessons 262*

A.2 Browser bookmarks 262

Goals for the case study 263 ■ Conventional bookmark

managers 263 ■ Modeling the bookmark collection 264

Simple means 268

A.3 Reflections on the case studies 268

references 269

index 275