

Part I Why Link Enterprise Data?

Semantic Web and the Linked Data Enterprise	3
Dean Allemang	
1 Social Data in the Enterprise	3
1.1 Causes	5
1.2 Technology Solutions	6
1.3 Localization and Globalization	9
2 The Linked Data Enterprise	10
2.1 Controlled Vocabularies	11
2.2 Prerequisites for Linked Data Vocabularies	18
3 Examples	20
3.1 Publishing	20
3.2 Government	21
4 Conclusions	22
References	22
The Role of Community-Driven Data Curation for Enterprises	25
Edward Curry, Andre Freitas, and Sean O’Riáin	
1 Introduction	25
2 The Business Need for Curated Data	26
3 Data Curation	28
3.1 How to Curate Data	28
4 Community-based Curated Enterprise Data	30
4.1 Internal Corporate Community	30
4.2 External Pre-competitive Communities	31
5 Case Study: Wikipedia - The World Largest Open Digital Curation Community	32
5.1 Social Organization	33
5.2 Artifacts, Tools and Processes	34
5.3 DBPedia - Community Curated Linked Open Data	35

6	Case Study: The New York Times - 100 Years of Expert Data	
	Curation	36
6.1	Data Curation	36
6.2	Publishing Curated Linked Data	37
7	Case Study: Thomson Reuters - Data Curation, a Core Business	
	Competency	38
7.1	Data Curation	39
8	Case Study: ChemSpider - Open Data Curation in the Global Chemistry Community	40
8.1	Community Objectives	41
8.2	Curation Approach & Types	41
9	Case Study: Protein Data Bank, Pre-competitive Bioinformatics ..	42
9.1	Serving the Community	42
9.2	Curation Approaches & Types	42
9.3	Observations	43
10	Case Study Learnings	44
10.1	Social Best Practices	44
10.2	Technical Best Practices	45
11	Conclusion	46
	References	46

Part II Approval and Support of Linked Data Projects

Preparing for a Linked Data Enterprise	51	
Bernadette Hyland		
1	Introduction	52
2	The Cost of Linked Data	52
2.1	The Cost of Services and Support	53
2.2	Education and Training	53
2.3	Infrastructure	54
3	Is your Organization Ready for Linked Data?	54
4	The Linked Data Initiative	57
5	A Decentralized Approach to Data Management	58
6	Being On the Web vs. In the Web	59
7	Leverage Vocabularies	60
8	A Simple Approach to Linked Data	61
9	Conclusions	62
9.1	Prepare for a Linked Data Enterprise	62
	References	63
Selling and Building Linked Data: Drive Value and Gain Momentum	65	
Kristen Harris		
1	The Data Burden	66
2	Driving Value Principles	67
3	Building a Team	70
4	Committing to Something Bigger	72

5	Putting it together	73
6	Conclusions	74
	References	76

Part III Techniques for Linking Enterprise Data

Enhancing Enterprise 2.0 Ecosystems Using Semantic Web and Linked Data Technologies: The SemSLATES Approach		79
Alexandre Passant, Philippe Laublet, John G. Breslin and Stefan Decker		
1	Introduction	80
2	Issues with Current Enterprise 2.0 Ecosystems	81
	2.1 Information Fragmentation and Heterogeneity of Data Formats	82
	2.2 Knowledge Capture and Re-use	83
	2.3 Tagging and Information Retrieval	83
3	SemSLATES: A Social and Semantic Middleware Approach for Enterprise 2.0	84
	3.1 The SemSLATES Architecture	85
	3.2 Ontologies for Enterprise 2.0	88
	3.3 Generating Semantic Annotations Through Software Add-ons	89
	3.4 Deploying Additional Services	90
4	Case-study: Enabling SemSLATES at EDF R&D	91
	4.1 Background	91
	4.2 Extending Popular Ontologies	92
	4.3 Automated SIOC-based Annotations	93
	4.4 Knowledge Capture Using UfoWiki	93
	4.5 Semantic Tagging Add-ons	95
	4.6 Additional Features of the Platform	96
5	Conclusion	99
	References	100
Linking XBRL Financial Data		103
Roberto García and Rosa Gil		
1	Introduction	103
	1.1 XBRL	106
	1.2 Related Work	108
2	Approach	109
	2.1 XSD2OWL Mapping	110
	2.2 XML2RDF Mapping	112
	2.3 Algorithm	113
3	Results	113
	3.1 Links to External Data	115
	3.2 Semantic Integration	118
4	Evaluation	119
	4.1 Use Case	122

5	Conclusions and Future Work	122
	References	124
Scalable Reasoning Techniques for Semantic Enterprise Data		
Reza B'Far		
1	Introduction	127
2	Survey of Reasoning Techniques	128
	2.1 Traditional Rule Engines	130
	2.2 Forward Chaining and the RETE algorithm	131
	2.3 Backward Chaining	132
3	Bayesian Networks	133
	3.1 Representing Probabilities within the Ontological Model	135
4	Unsupervised Reasoning	136
5	Semantic Reasoning	137
	5.1 Performance and Reasoning	139
	5.2 Applying Best-First Search (A* Search) to Semantic Reasoning	140
	5.3 High-level View of Distributed Reasoning	140
	5.4 Map-Reduce and Similar Techniques	141
	5.5 Performance and Ontology Engineering	143
6	Semantic Reasoning vs. Business Intelligence	143
7	Best Practices for Application Developers and System Integrators	144
8	Summary	146
	References	146
Reliable and Persistent Identification of Linked Data Elements		
David Wood		
1	Introduction	150
2	Metadata Before the World Wide Web	150
3	Metadata on the World Wide Web	154
4	Persistent URLs	159
5	Extending Persistent URLs for Web Resource Curation	160
6	Redirection of URL Fragments	164
7	Using Persistent URLs and Retrieved Metadata	164
8	Federations of PURL Servers	166
9	Conclusions and Further Work	170
	References	171
Part IV Success Stories		
Linked Data for Fighting Global Hunger: Experiences in setting standards for Agricultural Information Management		
Thomas Baker and Johannes Keizer		
1	Agricultural information and Semantic Web	177
2	Integrating access using Dublin Core metadata	180
3	AGROVOC and specialized domain ontologies	186

4	Networking, capacity development, and outreach	197
	References	201
Enterprise Linked Data as Core Business Infrastructure		203
Steve Harris and Tom Ilube and Mischa Tuffield		
1	Introduction	203
2	Motivations	204
3	Garlik's System Architectures	206
	3.1 DataPatrol	207
	3.2 QDOS	211
4	Schema Driven Software Deployment	215
5	Technology and the Need to Scale	216
	5.1 4store	216
	5.2 5store	217
6	Conclusions	218
7	Future Work	219
	References	219
Standardizing Legal Content with OWL and RDF		221
Constantine Hondros		
1	Introduction	221
	1.1 The problem domain	221
	1.2 Application of Semantic Web technologies	222
2	Toward a Common Legal Content Format	223
3	OWL Ontology	224
	3.1 Creating the ontology	224
	3.2 Domain Ontology Mapping	226
4	Content Architecture	227
	4.1 Modularized XHTML + RDFa for Textual Content	227
	4.2 RDF for Metadata, Relations and Classifications	228
5	Working with RDF in a Content Supply Chain	229
	5.1 The Open World Enigma	230
	5.2 Ensuring RDF Data Integrity	230
	5.3 Managing Fragmented Ontologies	232
	5.4 Managing Performance	232
	5.5 Using RDF with XSLT	233
6	Enabling Large-Scale Triple Production	234
	6.1 Experimental XSD generation	235
	6.2 RDFBeans	236
7	Conclusions	238
	References	239

A Role for Semantic Web Technologies in Patient Record Data Collection		241
Chimezie Ogbuji		
1	Introduction	242
2	Architectural Styles	243
	2.1 REST Architectural Style	243
	2.2 Service Oriented Architecture	244
3	Semantic Web Technologies	246
4	SemanticDB Concurrent Data Collection Workflow	247
	4.1 Requirements	248
	4.2 XML and RDF Content Management	249
	4.3 RESTful XSLT Services	250
	4.4 Declarative AJAX Framework	250
	4.5 Implementation	251
5	General Architectural Observations	257
6	Review of Service-oriented Metrics	257
7	Conclusions	260
	References	260
Use of Semantic Web technologies on the BBC Web Sites		263
Yves Raimond, Tom Scott, Silver Oliver, Patrick Sinclair and Michael Smethurst		
1	Introduction	263
	1.1 Linking microsites for cross-domain navigation	264
	1.2 Making data available to developers	264
	1.3 Making use of the wider Web	265
2	Programme support on the Web	265
	2.1 BBC Programmes	266
	2.2 The Programmes Ontology	266
	2.3 Web identifiers for broadcast radio and television sites	269
3	BBC Music	270
	3.1 BBC Music as Linked Data	271
	3.2 Web identifiers for BBC Music	271
	3.3 The Web as a content management system	272
	3.4 Using the BBC Programmes and the BBC Music Linked Data	272
4	BBC Wildlife Finder	274
	4.1 The Wildlife Ontology	274
	4.2 Web identifiers	276
	4.3 The Web as a Content Management System	278
	4.4 The importance of curation	278
5	Journalism	279
	5.1 Populating and using the ontology	280
	5.2 Future developments	281
6	Conclusion	282
	References	283

Contents	xvii
Glossary	285
Index	289