

Information Integration with Ontologies

Experiences from an Industrial Showcase

V. Alexiev
M. Breu
J. de Bruijn
D. Fensel
R. Lara
H. Lausen

 WILEY

Contents

<i>Foreword</i>	<i>ix</i>
<i>Acknowledgements</i>	<i>xi</i>
<i>List of Figures</i>	<i>xiii</i>
<i>1 Introduction</i>	<i>1</i>
<i>1.1 Finding a Way Out of the Dilemma</i>	<i>3</i>
<i>1.2 The Background to this Book</i>	<i>5</i>
<i>1.3 The Structure of the Book</i>	<i>5</i>
<i>1.3.1 Data modelling and ontologies</i>	<i>6</i>
<i>1.3.2 Information integration with relational databases and XML</i>	<i>6</i>
<i>1.3.3 The show case</i>	<i>6</i>
<i>1.3.4 Semantic information integration</i>	<i>7</i>
<i>1.3.5 Data source queries</i>	<i>7</i>
<i>1.3.6 Generating transformations</i>	<i>7</i>
<i>1.3.7 Best Practices and Methodologies</i>	<i>8</i>

2	<i>Data Modelling and Ontologies</i>	9
2.1	<i>The Information Integration Problem</i>	10
2.1.1	<i>How databases view the world</i>	11
2.1.2	<i>How ontologies view the world</i>	16
2.1.3	<i>Comparison</i>	20
2.2	<i>Semantic Information Management</i>	21
2.2.1	<i>Principles</i>	21
2.2.2	<i>The methodology</i>	24
2.3	<i>Conclusions</i>	26
3	<i>Information Integration with Relational Databases and XML</i>	27
3.1	<i>Introduction</i>	27
3.1.1	<i>Areas of data integration</i>	28
3.1.2	<i>Business drivers of data integration</i>	28
3.1.3	<i>Scope of this chapter</i>	31
3.2	<i>Relational Database Integration</i>	31
3.2.1	<i>Integration considerations</i>	31
3.2.2	<i>Integration approaches/degrees</i>	32
3.2.3	<i>Data centralization, sharing and federation</i>	33
3.2.4	<i>Integration characteristics</i>	33
3.3	<i>XML-based Integration</i>	42
3.3.1	<i>XML tools</i>	42
3.3.2	<i>XML and objects</i>	43
3.3.3	<i>XML and databases</i>	46
3.3.4	<i>XML transformations</i>	51
3.3.5	<i>XML, eCommerce and Web services</i>	53
3.4	<i>Conclusions</i>	53
3.4.1	<i>Summary</i>	54
3.4.2	<i>Variety in data integration</i>	54
4	<i>The Show Case</i>	55
4.1	<i>Data Sources</i>	56
4.2	<i>Identifying Overlaps between the Data Sources</i>	58
4.3	<i>Current Ways of Dealing with Heterogeneity</i>	59

5	<i>Semantic Information Integration</i>	63
5.1	<i>Approaches in Information Integration</i>	64
5.2	<i>Mapping Heterogeneous Data Sources</i>	65
5.2.1	<i>The Unicorn Workbench</i>	67
5.2.2	<i>Ontology construction and rationalization in the COG project</i>	75
5.3	<i>Other Methods and Tools</i>	81
5.3.1	<i>The MOMIS approach</i>	82
5.3.2	<i>InfoSleuth</i>	83
5.3.3	<i>OBSERVER</i>	86
5.3.4	<i>Ontology mapping in the KRAFT project</i>	89
5.3.5	<i>PROMPT</i>	91
5.3.6	<i>Chimæra</i>	94
5.3.7	<i>ONION</i>	96
5.3.8	<i>Other ontology merging methods</i>	99
5.4	<i>Comparison of the Methods</i>	100
5.4.1	<i>Comparison criteria</i>	100
5.4.2	<i>Comparing the methodologies for semantic schema integration</i>	104
5.5	<i>Conclusions and Future Work</i>	105
5.5.1	<i>Limitations of the Unicorn Workbench and future work</i>	106
6	<i>Data Source Queries</i>	107
6.1	<i>Querying Disparate Data Sources Using the Unicorn Workbench</i>	107
6.1.1	<i>Queries in the Unicorn Workbench</i>	108
6.1.2	<i>Transforming conceptual queries into database queries</i>	111
6.1.3	<i>Limitations of the current approach</i>	113
6.2	<i>Querying Disparate Data Sources</i>	113
6.2.1	<i>The querying architecture in the COG project</i>	114
6.2.2	<i>Querying in the COG showcase</i>	116
6.2.3	<i>Overcoming the limitations of the Unicorn Workbench</i>	120

6.3	<i>Related Work</i>	121
6.3.1	<i>Ontology query languages</i>	121
6.4	<i>Conclusions</i>	126
7	<i>Generating Transformations</i>	127
7.1	<i>Information Transformation in the COG Project</i>	128
7.1.1	<i>Generating transformations with the Unicorn Workbench</i>	129
7.1.2	<i>Automatic generation of transformations in the COG project</i>	133
7.2	<i>Other Information Transformation Approaches</i>	135
7.2.1	<i>Approaches that perform instance transformation</i>	138
7.2.2	<i>Approaches that do not perform instance transformation</i>	143
7.3	<i>Conclusions, Limitations and Extensions</i>	146
8	<i>Best Practices and Methodologies Employed</i>	149
8.1	<i>Best Practices</i>	149
8.1.1	<i>Selective mapping</i>	150
8.1.2	<i>Domain vs application modelling</i>	151
8.1.3	<i>Global-as-view vs local-as-view</i>	154
8.2	<i>Lessons Learned</i>	155
8.2.1	<i>Quality of global model depends on local models</i>	155
8.2.2	<i>Refinement of ontological concepts</i>	156
8.2.3	<i>Automation is hard to achieve in real-life situations</i>	157
8.2.4	<i>Queries vs transformations</i>	157
8.3	<i>Conclusions</i>	158
9	<i>Conclusion</i>	159
	<i>References</i>	163
	<i>Glossary</i>	173
	<i>Index</i>	177