

Contents

Abstract	iii
Resumé	v
Résumé	vii
1 Introduction	1
1 Semantic Web	1
2 Thesis Overview	5
2.1 Chapter 2: Towards Exploratory OLAP over Linked Open Data – A Case Study	6
2.2 Chapter 3: Processing Aggregate Queries in a Federa- tion of SPARQL Endpoints	7
2.3 Chapter 4: Efficient Support of Analytical SPARQL Queries in Federated Systems	9
2.4 Chapter 5: Optimizing Aggregate SPARQL Queries Us- ing Materialized RDF Views	10
3 Structure of the Thesis	12
2 Towards Exploratory OLAP over Linked Open Data – A Case Study	15
1 Introduction	16
2 A Movie Case Study	17
3 Source Discovery and Schema Building for Exploratory OLAP	23
3.1 Querying Knowledge Bases	24
3.2 Querying Data Management Platforms	26
3.3 Querying Semantic Web Search Engines	26
4 Conceptual Framework	27
5 Related Work	29
5.1 Semantic Web Data Warehousing	29
5.2 RDF Source Discovery	31
5.3 Indexing and Distributed Query Processing	32
6 Conclusions and Future Work	33

A	Prefixes Used in the Chapter	33
3	Processing Aggregate Queries in a Federation of SPARQL Endpoints	35
1	Introduction	36
2	Motivating Example and Preliminary Analysis	38
3	Related Work	40
4	Federated Processing of Aggregate Queries	41
5	Cost-Based Query Optimization	44
5.1	Query Optimizer	45
5.2	Cost Model	45
5.3	Estimating Constants	46
5.4	Result Size Estimation	47
6	Evaluation	49
6.1	Experimental Setup	49
6.2	Experimental Results	52
7	Conclusions and Future Work	54
4	Efficient Support of Analytical SPARQL Queries in Federated Systems	57
1	Introduction	58
2	Related Work	60
3	RDF Graph and Queries	61
4	Data Integration in LITE	62
4.1	Modeling Source and Target Schemas	63
4.2	Mapping Model	64
5	Query Rewriting and Optimization	67
5.1	Query Rewriting	67
5.2	Global and Local Optimization	70
5.3	RDF Entailment	72
6	Experimental Evaluation	73
6.1	Datasets, Setup, and Queries	74
6.2	Query Evaluation	75
7	Conclusion	77
5	Optimizing Aggregate SPARQL Queries Using Materialized RDF Views	79
1	Introduction	80
2	Related Work	81
3	RDF Graphs and Aggregate Queries	83
4	View Materialization in MARVEL	87
4.1	Creating Materialized RDF Views	87
4.2	Storing Materialized RDF Views	88
4.3	Data Cube Lattice	89

Contents

4.4	MARVEL Cost Model	90
4.5	RDF Entailment	92
4.6	MARVEL View Selection Algorithm	94
5	Query Rewriting in MARVEL	95
6	Evaluation	99
6.1	Datasets and Queries	100
6.2	Query Evaluation	104
7	Conclusion and Future Work	107
6	Conclusion	109
1	Summary of Results	109
2	Future Directions	111
	References	114