

# Contents

	<b>Preface</b> .....	<b>xv</b>
	<b>Acknowledgments</b> .....	<b>xix</b>
<b>1</b>	<b>Introduction</b> .....	<b>1</b>
<b>2</b>	<b>Data Graphs</b> .....	<b>5</b>
2.1	Models .....	5
2.1.1	Directed Edge-Labeled Graphs .....	6
2.1.2	Heterogeneous Graphs .....	8
2.1.3	Property Graphs .....	9
2.1.4	Graph Dataset .....	11
2.1.5	Other Graph Data Models .....	12
2.1.6	Graph Stores .....	13
2.2	Querying .....	13
2.2.1	Basic Graph Patterns .....	13
2.2.2	Complex Graph Patterns .....	16
2.2.3	Navigational Graph Patterns .....	19
2.2.4	Other Features .....	22
2.2.5	Query Interfaces .....	22
<b>3</b>	<b>Schema, Identity, and Context</b> .....	<b>25</b>
3.1	Schema .....	25
3.1.1	Semantic Schema .....	25
3.1.2	Validating Schema .....	27
3.1.3	Emergent Schema .....	32
3.2	Identity .....	35
3.2.1	Persistent Identifiers .....	35
3.2.2	External Identity Links .....	37
3.2.3	Datatypes .....	38
3.2.4	Lexicalization .....	38
3.2.5	Existential Nodes .....	39



3.3	Context .....	40
3.3.1	Direct Representation .....	40
3.3.2	Reification .....	41
3.3.3	Higher-Arity Representation .....	41
3.3.4	Annotations .....	42
3.3.5	Other Contextual Frameworks .....	44
<b>4</b>	<b>Deductive Knowledge .....</b>	<b>47</b>
4.1	Ontologies .....	48
4.1.1	Interpretations and Models .....	49
4.1.2	Ontology Features .....	51
4.1.3	Entailment .....	56
4.1.4	If-Then vs. If-and-Only-If Semantics .....	56
4.2	Reasoning .....	57
4.2.1	Rules .....	57
4.2.2	Description Logics .....	60
<b>5</b>	<b>Inductive Knowledge .....</b>	<b>67</b>
5.1	Graph Analytics .....	67
5.1.1	Techniques .....	69
5.1.2	Frameworks .....	70
5.1.3	Analytics on Data Graphs .....	74
5.1.4	Analytics with Queries .....	75
5.1.5	Analytics with Entailment .....	76
5.2	Knowledge Graph Embeddings .....	76
5.2.1	Tensor-Based Models .....	78
5.2.2	Language Models .....	88
5.2.3	Entailment-Aware Models .....	90
5.3	Graph Neural Networks .....	91
5.3.1	Recursive Graph Neural Networks .....	91
5.3.2	Non-Recursive Graph Neural Networks .....	95
5.4	Symbolic Learning .....	96
5.4.1	Rule Mining .....	97
5.4.2	Axiom Mining .....	100
5.4.3	Hypothesis Mining .....	102



<b>6</b>	<b>Creation and Enrichment</b> .....	<b>105</b>
6.1	Human Collaboration .....	105
6.2	Text Sources .....	106
6.2.1	Pre-Processing .....	106
6.2.2	Named Entity Recognition (NER) .....	107
6.2.3	Entity Linking (EL) .....	107
6.2.4	Relation Extraction (RE) .....	108
6.2.5	Joint Tasks .....	109
6.3	Markup Sources .....	109
6.3.1	Wrapper-Based Extraction .....	110
6.3.2	Web Table Extraction .....	111
6.3.3	Deep Web Crawling .....	111
6.4	Structured Sources .....	111
6.4.1	Mapping from Tables .....	112
6.4.2	Mapping from Trees .....	114
6.4.3	Mapping from Other Knowledge Graphs .....	114
6.5	Schema/Ontology Creation .....	115
6.5.1	Ontology Engineering .....	115
6.5.2	Ontology Learning .....	116
<b>7</b>	<b>Quality Assessment</b> .....	<b>119</b>
7.1	Accuracy .....	119
7.1.1	Syntactic Accuracy .....	119
7.1.2	Semantic Accuracy .....	120
7.1.3	Timeliness .....	120
7.2	Coverage .....	121
7.2.1	Completeness .....	121
7.2.2	Representativeness .....	121
7.3	Coherency .....	122
7.3.1	Consistency .....	122
7.3.2	Validity .....	123
7.4	Succinctness .....	123
7.4.1	Conciseness .....	123
7.4.2	Representational Conciseness .....	124
7.4.3	Understandability .....	124
7.5	Other Quality Dimensions .....	125



<b>8</b>	<b>Refinement</b> .....	<b>127</b>
8.1	Completion .....	127
8.1.1	General Link Prediction .....	127
8.1.2	Type-Link Prediction .....	128
8.1.3	Identity-Link Prediction .....	128
8.2	Correction .....	129
8.2.1	Fact Validation .....	129
8.2.2	Inconsistency Repairs .....	131
8.3	Other Refinement Tasks .....	131
<b>9</b>	<b>Publication</b> .....	<b>133</b>
9.1	Best Practices .....	133
9.1.1	FAIR Principles .....	133
9.1.2	Linked Data Principles .....	135
9.2	Access Protocols .....	137
9.2.1	Dumps .....	137
9.2.2	Node Lookups .....	138
9.2.3	Edge Patterns .....	138
9.2.4	(Complex) Graph Patterns .....	139
9.2.5	Other Protocols .....	140
9.3	Usage Control .....	140
9.3.1	Licensing .....	140
9.3.2	Usage Policies .....	141
9.3.3	Encryption .....	142
9.3.4	Anonymization .....	143
<b>10</b>	<b>Knowledge Graphs in Practice</b> .....	<b>145</b>
10.1	Open Knowledge Graphs .....	145
10.1.1	DBpedia .....	145
10.1.2	Yet Another Great Ontology .....	146
10.1.3	Freebase .....	147
10.1.4	Wikidata .....	147
10.1.5	Other Open Cross-Domain Knowledge Graphs .....	148
10.1.6	Domain-Specific Open Knowledge Graphs .....	148
10.2	Enterprise Knowledge Graphs .....	149
10.2.1	Web Search .....	149
10.2.2	Commerce .....	150

10.2.3	Social Networks .....	150
10.2.4	Finance .....	151
10.2.5	Other Industries .....	151
<b>11</b>	<b>Conclusions .....</b>	<b>153</b>
<b>A</b>	<b>Background .....</b>	<b>157</b>
A.1	Historical Perspective .....	157
A.2	“Knowledge Graphs:” Pre-2012 .....	158
A.3	“Knowledge Graphs:” 2012 Onward .....	161
	<b>Bibliography .....</b>	<b>165</b>
	<b>Authors’ Biographies .....</b>	<b>229</b>

Graphs, Networks (in their various forms), and the Semantic Web. As the discussion continued, a consensus began to emerge. Knowledge Graphs, as a topic, involved a novel confluence of techniques originating from previously disparate scientific communities, each the harbinger of developing novel graph-based techniques for better integrating and extracting value from diverse knowledge sources at large scale.

As a follow-up to the summit, the attendees agreed that in order to foster this growing view of Knowledge Graphs, there was a need for a manuscript that would serve as a general introduction to the area. This manuscript would:

- introduce knowledge graphs and the value of abstracting data as graphs;
- survey the historical context of knowledge graphs and the key initiatives leading to their popularization;
- draw together disparate views of knowledge graphs into a unifying definition;
- provide an introduction to the key techniques that knowledge graphs enable, relating to querying, validation, reasoning, learning, refinement, enrichment, quality assessment, and more besides;
- describe how knowledge graphs are used in practice, surveying the companies using knowledge graphs, the applications they are used for, the open knowledge graphs that have been published, etc.; and
- delineate future research directions for knowledge graphs.

The manuscript would then serve as an introductory text for students, practitioners and researchers new to the area, helping to form a consensus of views of what is a knowledge graph, laying the foundations for future developments.