

# Contents

<b>1. Introduction</b>	<b>1</b>
1.1. Motivation . . . . .	1
1.2. Main Challenge . . . . .	2
1.3. Proposed Solution . . . . .	3
1.4. Hypothesis . . . . .	3
1.4.1. What representations of a versioned dataset allow for better compression, more efficient indexing, and/or more efficient updates? . . . . .	4
1.4.2. How should queries be rewritten to be compatible with a versioned dataset? . . . . .	4
1.4.3. How large is the performance cost of implementing a versioned dataset and query system? . . . . .	5
1.4.4. Which of the representation alternatives will have the most well-rounded performance? . . . . .	5
1.5. Objectives . . . . .	5
1.5.1. General Objectives . . . . .	5
1.5.2. Specific Objectives . . . . .	6
<b>2. Related Work</b>	<b>7</b>
2.1. Semantic Web . . . . .	7
2.1.1. RDF . . . . .	7
2.1.2. SPARQL . . . . .	8
2.1.3. Linked Data . . . . .	9
2.2. Wikidata . . . . .	11
2.3. Versioning of RDF . . . . .	12
<b>3. Preliminaries</b>	<b>15</b>
3.1. RDF . . . . .	15
3.1.1. RDF terms . . . . .	15
3.1.2. RDF triple . . . . .	16
3.1.3. RDF graph and dataset . . . . .	16
3.2. SPARQL . . . . .	16
3.2.1. Basic Concepts . . . . .	16
3.2.2. Basic Graph Pattern . . . . .	17
3.2.3. SPARQL query . . . . .	17
3.2.4. Solution mapping . . . . .	18
3.2.5. SPARQL 1.1 . . . . .	20

<b>4. Native RDF Versioning</b>	<b>22</b>
4.1. Preliminary Concepts . . . . .	22
4.1.1. Versioned Graph . . . . .	22
4.1.2. Graph Deltas . . . . .	23
4.1.3. Temporal Query . . . . .	23
4.1.3.1. Single Version query . . . . .	23
4.1.3.2. Version Delta query . . . . .	23
4.2. Running Example . . . . .	23
4.3. Complete Versions . . . . .	25
4.3.1. Construction details . . . . .	25
4.3.1.1. Dataset construction . . . . .	25
4.3.1.2. Query conversion . . . . .	25
4.3.2. Expected advantages . . . . .	25
4.3.3. Expected disadvantages . . . . .	25
4.4. Interval-based graphs . . . . .	27
4.4.1. Construction Details . . . . .	27
4.4.1.1. Dataset construction . . . . .	27
4.4.1.2. Query conversion . . . . .	28
4.4.2. Expected advantages . . . . .	28
4.4.3. Expected disadvantages . . . . .	28
4.5. Absolute deltas . . . . .	30
4.5.1. Construction details . . . . .	30
4.5.1.1. Dataset construction . . . . .	30
4.5.1.2. Query conversion . . . . .	31
4.5.2. Expected advantages . . . . .	31
4.5.3. Expected disadvantages . . . . .	34
4.6. Sequential deltas . . . . .	34
4.6.1. Construction details . . . . .	34
4.6.1.1. Dataset construction . . . . .	34
4.6.1.2. Query conversion . . . . .	35
4.6.2. Expected advantages . . . . .	37
4.6.3. Expected disadvantages . . . . .	37
4.7. Complex queries . . . . .	37
4.8. Implementation . . . . .	38
<b>5. Evaluation</b>	<b>39</b>
5.1. Experimental goals . . . . .	39
5.2. Experimental settings . . . . .	39
5.2.1. System specifications . . . . .	39
5.2.2. Dataset and representations . . . . .	40
5.2.3. Experiments . . . . .	40
5.3. Experimental results . . . . .	41
5.3.1. Dataset size . . . . .	41
5.3.2. Storage size . . . . .	41
5.3.3. Construction times . . . . .	42
5.3.4. Query times . . . . .	44
5.3.4.1. Single Version queries . . . . .	44

5.3.4.2. Version Delta queries . . . . .	46
5.4. Synthesis . . . . .	49
<b>6. Conclusions</b>	<b>50</b>
6.1. Summary . . . . .	50
6.2. Scope and limitations . . . . .	50
6.2.1. Query engine . . . . .	50
6.2.2. Query set . . . . .	51
6.2.3. Dataset . . . . .	51
6.2.4. Versioning representations . . . . .	51
6.2.5. Property paths . . . . .	51
6.3. Future Work . . . . .	51
<b>Bibliography</b>	<b>53</b>