

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Knowledge Engineering . . . . .	3
1.2	Conceptual Knowledge Processing . . . . .	5
1.3	Web Mining . . . . .	6
1.4	Overview . . . . .	8
<b>2</b>	<b>Foundations</b>	<b>11</b>
2.1	Ordered Sets and Lattices . . . . .	11
2.1.1	Ordered Sets . . . . .	11
2.1.2	Lattices as Ordered Sets . . . . .	12
2.2	Formal Concept Analysis . . . . .	13
2.2.1	Formal Context . . . . .	14
2.2.2	Formal Concept . . . . .	16
2.2.3	Gaining Sets of Concepts from Contexts . . . . .	16
2.2.4	Conceptual Hierarchy . . . . .	17
2.2.5	Concept Lattice . . . . .	18
2.2.6	Line diagram of a Concept Lattice . . . . .	19
2.2.7	Dealing with Complexity . . . . .	20
2.3	Summary . . . . .	21
<b>3</b>	<b>Web Search</b>	<b>23</b>
3.1	Web Objects . . . . .	23
3.2	Web Search . . . . .	24
3.2.1	Search Query . . . . .	25
3.2.2	Search Result List . . . . .	26
3.2.3	Formalization of Web Search Results . . . . .	28
3.3	Web Information Retrieval . . . . .	29
3.3.1	Web Information Retrieval System . . . . .	29
3.3.2	Relevance . . . . .	30

3.3.3	Feature Engineering . . . . .	30
3.4	Web Search Engines . . . . .	32
3.4.1	Major Web Search Engines . . . . .	32
3.4.2	Catalog vs. Crawler . . . . .	36
3.4.3	Generations . . . . .	36
3.5	Summary . . . . .	38
<b>4</b>	<b>Quality of Search Results</b>	<b>39</b>
4.1	Snippets . . . . .	39
4.1.1	Concordances and Keywords in Context . . . . .	40
4.1.2	Types of Snippets . . . . .	41
4.2	Ranking . . . . .	42
4.2.1	Relevance . . . . .	42
4.2.2	Precision vs. Recall . . . . .	42
4.2.3	HITS . . . . .	43
4.2.4	PageRank . . . . .	46
4.2.5	Other Ranking Criteria . . . . .	47
4.3	Search Engine Overlap & Size of the Web . . . . .	49
4.4	Commercial Interests . . . . .	50
4.5	Summary . . . . .	51
<b>5</b>	<b>FooCA: A Conceptual Search Engine</b>	<b>53</b>
5.1	Background . . . . .	53
5.2	Clustering and FCA . . . . .	54
5.3	Retrieval via a standard Web Search Engine . . . . .	55
5.4	From Ranked Lists to Conceptual Structures . . . . .	56
5.5	Architecture . . . . .	56
5.6	Feature Extraction . . . . .	58
5.7	From a Web Search Engine Retrieval to a Context . . . . .	59
5.8	Representing the Context in a Cross Table . . . . .	59
5.9	Search Preferences and Strategies . . . . .	61
5.9.1	Choice of Search Engine . . . . .	61
5.9.2	Interval of the Search Result Retrieval . . . . .	61
5.9.3	Language Restriction . . . . .	61
5.9.4	Removal of Stop Words . . . . .	62
5.9.5	Stemming . . . . .	62

5.9.6	Clarification of the Context . . . . .	62
5.9.7	User-based Query Refinement . . . . .	62
5.9.8	Limiting by an Object Count for Attributes . . . . .	64
5.9.9	Minimum Attribute Length . . . . .	64
5.9.10	Attribute Ranking . . . . .	64
5.9.11	Exporting the Context . . . . .	65
5.10	Web-based Visualization of the Concept Lattice . . . . .	66
5.11	Summary . . . . .	69
<b>6</b>	<b>Examples and Evaluation</b>	<b>71</b>
6.1	Use-Case Scenario . . . . .	71
6.1.1	Stepwise Context Size Reduction via a Cross Table . . . . .	72
6.1.2	Stepwise Context Refinement via a Cross Table . . . . .	73
6.1.3	Visualization of the Concept Lattice using a Line Diagram . . . . .	75
6.2	Examples of FooCA generated Concept Lattices . . . . .	77
6.3	Complexity . . . . .	83
6.4	Usability . . . . .	86
6.5	Summary . . . . .	87
<b>7</b>	<b>Conclusions</b>	<b>89</b>
7.1	Discussion . . . . .	89
7.2	Related Work . . . . .	92
7.3	Future Work . . . . .	93
7.4	Outlook . . . . .	94
	<b>Appendix</b>	<b>97</b>
	<b>Bibliography</b>	<b>101</b>