

Inhalt

| | |
|---|-----------|
| Vorwort | 3 |
| 1 Einführung | 14 |
| 2 Semantic Web Standards und Formate | 16 |
| 2.1 W3C- Standards | 17 |
| 2.1.1 Resource Description Framework (RDF) | 17 |
| 2.1.2 RDF Schema (RDFS) | 22 |
| 2.1.3 Web Ontology Language (OWL) | 24 |
| 2.1.4 SPARQL Protocol And RDF Query Language (SPARQL) | 24 |
| 2.2 W3C-Empfehlungen | 25 |
| 2.2.1 Rule Interchange Format (RIF) | 26 |
| 2.2.2 Semantic Web Rule Language (SWRL) | 27 |
| 2.2.3 Microdata, Microformat und RDFa | 27 |
| 2.2.4 Simple Knowledge Organisation System (SKOS) | 30 |
| 2.2.5 RDB2RDF | 30 |
| 2.3 Weitere semantische Beschreibungssprachen | 30 |
| 2.3.1 F-Logic | 30 |
| 2.3.2 Prolog | 32 |
| 3 Formale Semantik von Ontologiesprachen | 33 |
| 3.1 Formale Semantik von RDF(S) | 33 |
| 3.2 Schlussfolgerungen mit RDF(S) | 39 |
| 3.3 Formale Semantik von OWL | 49 |

| | | |
|----------|--|-----------|
| 3.4 | Schlussfolgerungen mit OWL | 56 |
| 3.5 | Open / Closed World Assumption | 64 |
| 4 | Linked Open Data (LOD) | 65 |
| 4.1 | LOD – Datensätze | 65 |
| 4.1.1 | DBpedia | 67 |
| 4.1.2 | Freebase | 67 |
| 4.1.3 | WordNet | 67 |
| 4.1.4 | YAGO | 68 |
| 4.1.5 | Friend of a friend (FOAF) | 68 |
| 4.2 | LOD – Registries und Datenmanagementsysteme | 69 |
| 4.2.1 | Open Metadata Registry (OMR) | 69 |
| 4.2.2 | Linked Open Vocabularies (LOV) | 70 |
| 4.2.3 | Comprehensive Knowledge Archive Network (CKAN) | 70 |
| 4.3 | LOD2 | 71 |
| 4.4 | LOD – Beispiele aus der Praxis | 71 |
| 4.4.1 | LOD – Energiedaten | 72 |
| 4.4.2 | LOD – Regierungsdaten | 72 |
| 4.4.3 | LOD – Bibliotheksdaten | 73 |
| 5 | Prozessschritte der semantischen Suche | 74 |
| 5.1 | Beschaffung der Datenbasis | 74 |
| 5.2 | Aufbereitung der Datenbasis | 75 |
| 5.3 | Erstellung der Suchanfrage | 75 |

| | | |
|----------|---|-----------|
| 5.4 | Analyse und Aufbereitung der Suchanfrage | 75 |
| 5.5 | Durchführung der Suchanfrage | 76 |
| 5.6 | Analyse und Aufbereitung der Suchergebnisse | 76 |
| 5.7 | Präsentation der Suchergebnisse | 76 |
| 5.8 | Verfeinerung der Suchanfrage | 77 |
| 6 | Methoden und Techniken der semantischen Suche | 78 |
| 6.1 | Methoden zur Gewinnung der Datenbasis | 78 |
| 6.1.1 | Web Crawling | 78 |
| 6.1.2 | Focused Web Crawling | 80 |
| 6.1.3 | Semantic Data Crawling | 80 |
| 6.1.4 | Metasuche | 81 |
| 6.2 | Methoden zur Aufbereitung der Datenbasis | 81 |
| 6.2.1 | Erstellung eines semantischen Graphen auf Basis der Datenstruktur | 81 |
| 6.2.2 | Interpretation der Inhalte | 82 |
| 6.2.3 | Ableitung von zusätzlichem Wissen | 83 |
| 6.3 | Suchmethoden | 84 |
| 6.3.1 | Keyword Search | 84 |
| 6.3.2 | Natural Language Search | 85 |
| 6.3.3 | Exploratory Search | 85 |
| 6.4 | Methoden zur Interpretation der Suchanfrage | 86 |
| 6.4.1 | Concept Detection | 86 |
| 6.4.2 | Übersetzung von Keywords in strukturierte Suchanfragen | 87 |

| | | |
|----------|---|------------|
| 6.4.3 | Übersetzung von natürlicher Sprache in strukturierte Suchanfragen | 88 |
| 6.4.4 | Iterative Query Construction | 89 |
| 6.5 | Matching-Methoden | 89 |
| 6.5.1 | Term-basiertes Matching | 90 |
| 6.5.2 | Graph-basiertes Matching | 90 |
| 6.5.3 | Reasoning-basiertes Matching | 91 |
| 6.6 | Ranking-Methoden | 92 |
| 6.6.1 | Hyperlinkanalyse | 93 |
| 6.6.2 | Analyse semantischer Verbindungen | 95 |
| 6.6.3 | Analyse der Keyword-Nähe | 95 |
| 6.6.4 | Analyse der Keyword-Häufigkeit | 96 |
| 6.7 | Methoden zur Darstellung der Suchergebnisse | 97 |
| 6.7.1 | Ergebnislisten | 98 |
| 6.7.2 | Graphen | 99 |
| 6.7.3 | Topic Maps | 100 |
| 6.7.4 | Geografische Karten | 102 |
| 7 | Komponenten für die Entwicklung Semantischer Suchsysteme | 104 |
| 7.1 | Web Crawling Tools | 106 |
| 7.1.1 | CRAWLJAX | 106 |
| 7.1.2 | crawler4j | 107 |
| 7.1.3 | ALVIS | 107 |
| 7.1.4 | Apache Nutch | 108 |

| | | |
|-------|--|-----|
| 7.1.5 | LDSpider | 108 |
| 7.1.6 | Virtuoso Sponger | 108 |
| 7.2 | Information Retrieval Libraries | 109 |
| 7.2.1 | Dragon Toolkit | 109 |
| 7.2.2 | Apache Lucene | 109 |
| 7.2.3 | Xapian | 110 |
| 7.3 | Entity Extraction Tools | 110 |
| 7.3.1 | Apache Unstructured Information Management Architecture (UIMA) | 110 |
| 7.3.2 | OpenCalais | 111 |
| 7.3.3 | AlchemyAPI | 112 |
| 7.3.4 | Stanford Named Entity Recognizer (NER) | 112 |
| 7.3.5 | General Architecture for Text Engineering (GATE) | 112 |
| 7.4 | Reasoner | 113 |
| 7.4.1 | FaCT++ | 113 |
| 7.4.2 | HermiT | 114 |
| 7.4.3 | Pellet | 114 |
| 7.4.4 | KAON2 | 114 |
| 7.5 | Semantic Matching Frameworks und Engines | 115 |
| 7.5.1 | S-Match | 115 |
| 7.5.2 | MatchIT | 115 |
| 7.6 | Semantic Vizualisation Tools und Frameworks | 116 |
| 7.6.1 | Recline.js | 116 |

| | | |
|----------|---|------------|
| 7.6.2 | D3.js | 116 |
| 7.6.3 | BubbleTree | 117 |
| 7.6.4 | jOWL | 117 |
| 7.6.5 | Thinkmap | 117 |
| 7.6.6 | Callimachus | 117 |
| 7.7 | Semantic Web Application Frameworks | 118 |
| 7.7.1 | Apache Jena | 118 |
| 7.7.2 | Sesame | 119 |
| 7.7.3 | CubicWeb | 121 |
| 7.7.4 | Semantic Information Logistics Architecture (SMILA) | 122 |
| 7.7.5 | Sphinx | 124 |
| 7.8 | Einordnung der Werkzeuge | 125 |
| 8 | Semantische Suchmaschinen | 128 |
| 8.1 | Bestehende semantische Online-Suchsysteme | 129 |
| 8.2 | Detaillierte Betrachtung ausgewählter Systeme | 131 |
| 8.2.1 | Lexxe | 132 |
| 8.2.2 | SHOE Language und SHOE Search | 134 |
| 8.2.3 | Swoogle | 136 |
| 9 | Zusammenfassung | 139 |
| | Literaturverzeichnis | 141 |