



Information Retrieval on the Semantic Web - Does it exist?

aposdle – New ways ...

... to work, learn and collaborate!

Peter Scheir, Viktoria Pammer, Stefanie N. Lindstaedt

Situation

- **Good amount of retrieval approaches associated with the Semantic Web**
- **Which of them qualify as IR for the Semantic Web?**

Criteria for information retrieval on the Semantic Web

- **The system operates on the Semantic Web**
- **The system is based on technology for the Semantic Web**
- **The system performs information retrieval (not data retrieval)**

Semantic Web vs. Semantic Desktop

- **Most approaches target Semantic Desktop**
 - Operate in desktop environment
 - Incubation of Semantic Web technology on desktop

Ontology-driven information retrieval vs. information retrieval for the Semantic Web

- **Not the same thing**
- **Many Semantic Web approaches use ontologies to increase retrieval performance**
 - Not all do. E.g. Statistical processing of RDF triples
- **It does not have to be IR in SW if an ontology is used**

Data Retrieval vs. Information Retrieval

- **Some of current approaches are based on query languages for the Semantic Web**
 - As e.g. SPARQL
- **All results are equally(!) relevant**

Surveyed systems

- **If system called “for the Semantic Web” by authors**
- **If systems uses technology from the Semantic Web**
- **Two types of systems surved**
 - Systems or models that search for information in the form of documents or ontological elements
 - Approaches that search for ontologies (being a special type of information on the Semantic Web)

Search for documents or ontological elements

- 20 systems surveyed
- Classified by:
 - **(Semantic) Web / (Semantic) Desktop**
 - **Concept based query / Term based query**
 - **Knowledge-based system / Information retrieval system**
 - **Data retrieval / Information retrieval**
 - **Knowledge representation used**

Search for ontologies

- 6 systems surveyed
- Classified by:
 - **Input type (free text, keyword, formal element, formal structure)**
 - **Crawler (Yes, No)**
 - **Storage of ontologies (Yes, No)**
 - **Indexing of ontologies (Yes, No)**
 - **Supported knowledge representations**
 - **Does an API exist (Yes, No)**
 - **Available online (Yes, No)**

Conclusions (document search)

- **None of the surveyed systems for search for documents and ontological concepts did fulfil all three characteristics**
 - operate in Semantic Web
 - Semantic Web technology (not “just” ontology based IR)
 - information retrieval
- **Most operate on desktop**
- **Not all use technology for the Semantic Web**
- **Some do data retrieval**

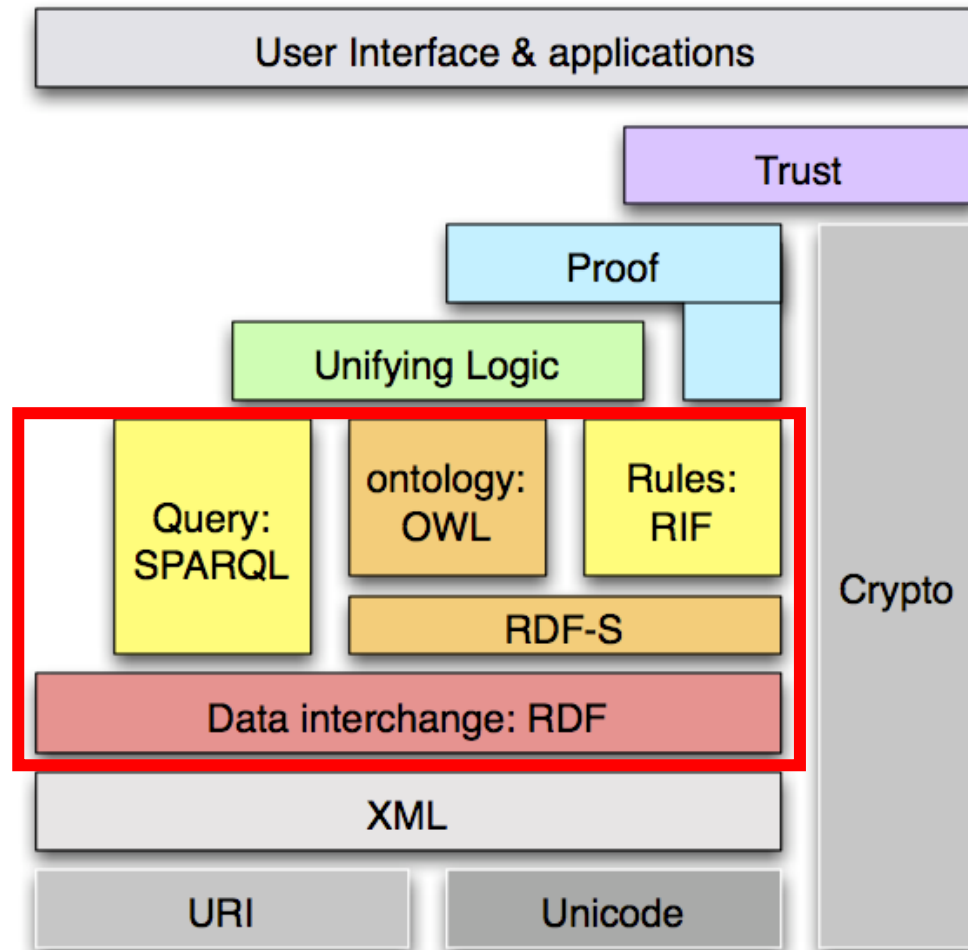
Conclusions (ontology search)

- **Not many ontology search engines**
- **But adhere the three criteria defined**

Conclusions

- **All systems operate on lower layers of Semantic Web**
- **Proof and trust not addresses in systems**

Conclusions



[Berners-Lee 2006]

Thank you for your attention!

- Questions / Comments?
- pscheir@know-center.at
- <http://www.aposdle.org/>



Our questions

- **Did you find the definition of IR on the SW fitting?**
- **Did you find the characteristics for surveying systems appropriate?**

Criteria for information retrieval on the Semantic Web

- **The system operates on the Semantic Web**
- **The system is based on technology for the Semantic Web**
- **The system performs information retrieval (not data retrieval)**

Search for documents or ontological elements

- 20 systems surveyed
- Classified by:
 - **(Semantic) Web / (Semantic) Desktop**
 - **Concept based query / Term based query**
 - **Knowledge-based system / Information retrieval system**
 - **Data retrieval / Information retrieval**
 - **Knowledge representation used**

Search for ontologies

- 6 systems surveyed
- Classified by:
 - **Input type (free text, keyword, formal element, formal structure)**
 - **Crawler (Yes, No)**
 - **Storage of ontologies (Yes, No)**
 - **Indexing of ontologies (Yes, No)**
 - **Supported knowledge representations**
 - **Does an API exist (Yes, No)**
 - **Available online (Yes, No)**