Connecting the Dots of Linked Data of Resource Collections

Thorsten Liebig | liebig@derivo.de
• SME, est. 2010
• Design and development of semantic software solutions
• Expertise: Smart Data, Big Data and Semantic Technologies
• OWLlink, OWL API
• Clients: intelligence authorities machine engineering publisher
RDF Graphs

Pro:
• simplest possible structure
• easy to produce, exchange, consume, ...
• straightforward query language

Con:
• everything has to be a triple
  (reification is painful)
RDF Graphs

Pro:
• simplest possible structure
• easy to produce, exchange, consume, ...
• straightforward query language

Con:
• everything has to be a triple (reification is painful)
• mix-up of data and schema
RDF Graphs

Pro:
• simplest possible structure
• easy to produce, exchange, consume, ...
• straightforward query language

Con:
• everything has to be a triple (reification is painful)
• mix-up of data and schema
What’s in a RDF Graph?

- Sandra:Female
  - rdf:type: Female
  - rdfs:subClassOf: Person
  - rdfs:subClassOf: Female
- Jack
  - has-sibling: Mike
  - rdf:type: Male
  - rdfs:subClassOf: Person
- Mike
  - age: 28
  - rdf:type: Male
  - rdfs:subClassOf: Person
What’s in a RDF Graph?

**Schema** (TBox)
- vocabulary of domain
- vocabulary conform data
- knowledge about domain

**Data** (ABox)
- data of the domain
- object of queries
- contains inferred facts
- larger than schema and subject to updates

---

**Diagram:**

- Person
- Female
- Male
- rdfs:subClassOf
- relative
- is-sibling
- has-son
- age [range: xsd:int]

- Mike:Male
- Sandra:Female
- Jack
- has-son
- has-nephew
- is-sibling
- symmetric: is-sibling
- is-sibling $\oplus$ has-son $\Rightarrow$ has-nephew
How to Query and Visualize a RDF Graph?

SPARQL:
SELECT ?x ?y WHERE {
  ?x a Female.
  ?x relative ?y.
}

http://yasgui.org
How to Query and Visualize a RDF Graph?

SPARQL:
SELECT ?x ?y WHERE {
  ?x a Female .
  ?x relative ?y .
}

Mike:Male
Sandra:Female
Jack
Sandra is-sibling Mike
Sandra has-nephew Mike
Mike has-son Jack
Mike age 28
Sandra age 28
Person rdfs:subClassOf Female Male
is-sibling relative
has-son
age [range: xsd:int]

http://www.irisa.fr/LIS/ferre/sparklis
Faceted Search

- http://www.cs.ox.ac.uk/isg/tools/SemFacet/
- well known to users
- allows logic operators (and/or)
- focus on one result variable (when compared to SPARQL)
- overview missing
Network Visualizations
SemSpect

Interactive, data driven visualization and analysis via:

• grouping of nodes
• aggregation of relations
• selective exploration
• on-demand details
• sophisticated filtering
• http://www.semspect.de
Wrap-Up

Lesson learned
- business critical data is not always perfect
- semantic modelling is not for free but pays off
- don’t get confused by the triple heap

Key factors for effective querying and exploring
- user guidance wrt. schema and query syntax
- scalability of visualization paradigm
- user orientation and result interpretation
- data-driven exploration options