



semdata.org

# Semantic Data Management Initiative

Introduction, SemData@ESWC, Crete

<http://www.semadata.org>

Atanas Kiryakov, Ontotext

---

May, 2010

# Semantic Technology

---

- Over the last ten years the **enthusiasm** about SemTech have been growing steadily
- However, they often appear to be **immature** and tend to be too expensive or **risky for deployment** in real business
- **Solid data management layer** is important to everyone in the semantic technology ecosystem
- Shaping it requires a **strong community** and a critical mass



# The SemData Initiative

---

- SemData is about **semantic databases and related storage and retrieval infrastructure**
- The SemData initiative brings **series of events** aim to facilitate the development and adoption of semantic data management concepts, tools, and practices
- SemData aims to **foster the development of a community**, which can bring about the tipping point where semantic technologies become successful and realize their full potential



# Semantic Repositories

---

- Semantic repositories combine characteristics of:
  - database management systems (**DBMS**) and
  - **inference engines**
- The **web servers of Web 3.0**
  - They are able to hold, interpret, and serve requests from multiple users to massive amounts of data
- Semantic repositories are still in the initial phase of **rapid development**
  - Since 2004, each couple of years, the engines have been getting an order of magnitude faster and more scalable



# Semantic Repositories = Track-laying machines

---

- Each new development allows loading more data, dealing with more comprehensive schemata and ontologies, and answering more complex queries in less time
- As in **mountain climbing**, each new achievement opens new opportunities and challenges
- Semantic repositories can also be seen as **track-laying** machines, which extend the reach of the data railways, step by step, changing the data-economy of entire domains and areas, by allowing more and more complex data to be handled at lower cost



# Semantic Repositories = Track-laying machines (2)

---



Property of MSCUA, University of Washington Libraries. Photo Coll 516



# Semantic Repositories = Track-laying machines (3)

---



# Semantic Repositories

---

We build upon **lightweight semantics** that is easy to understand, deploy, and manage

For instance, think of ontologies as database schemata with simple interpretation rules. Plenty of obvious (but useful) implicit facts can be inferred and match queries right away

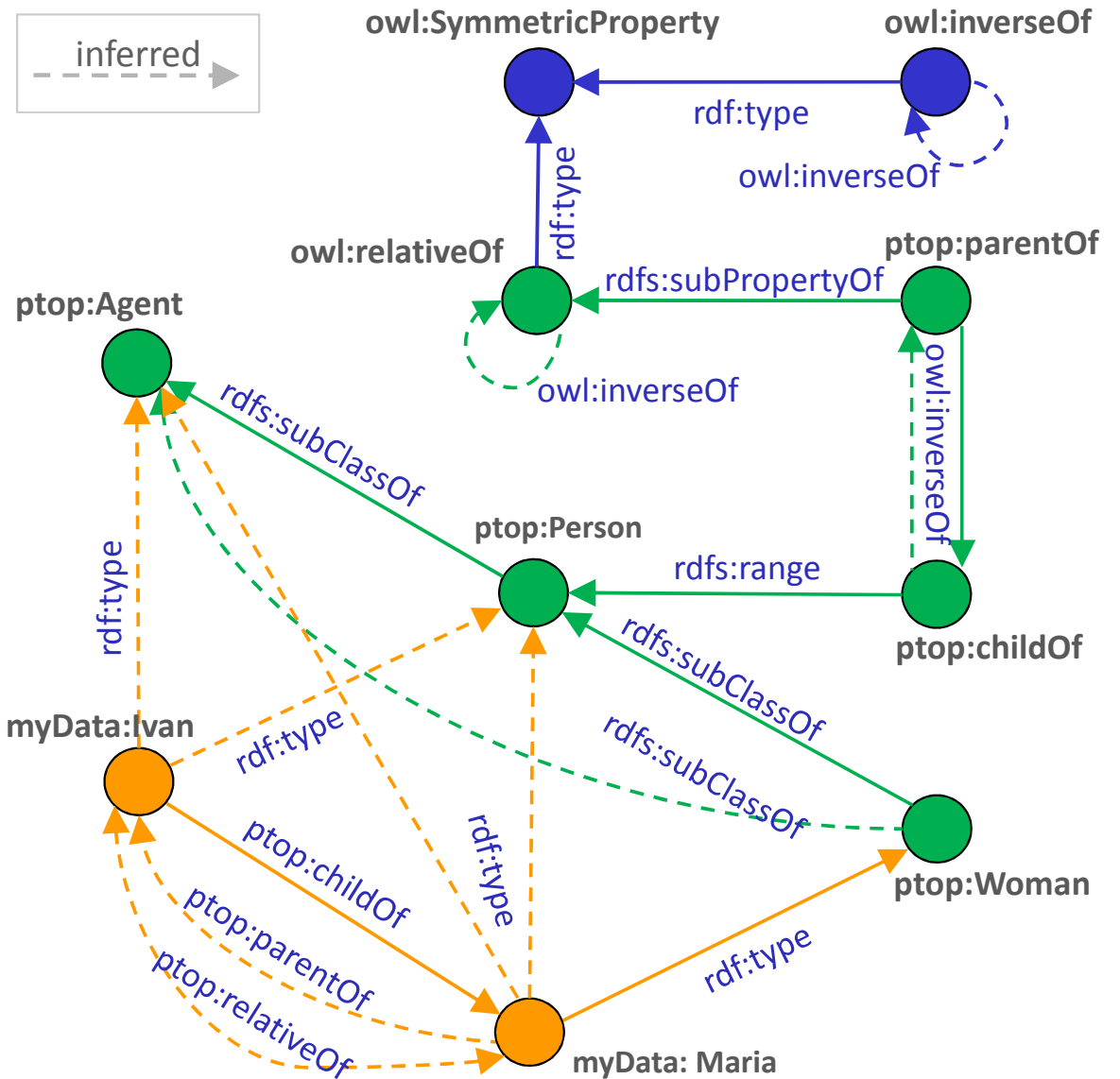




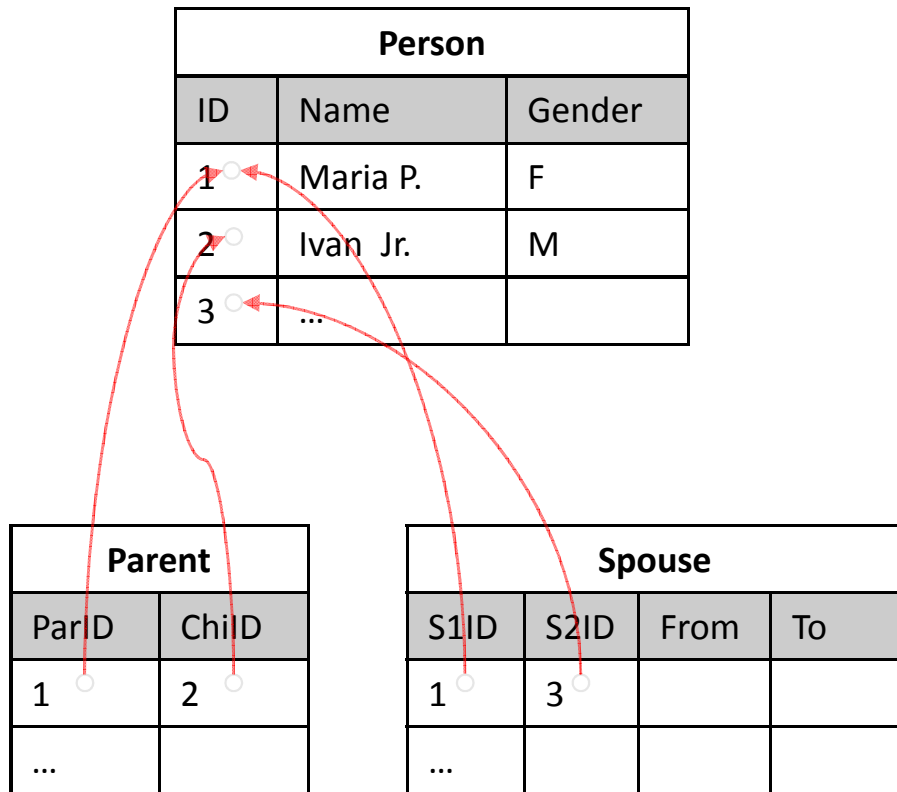
# Get more facts – Match more queries

The database will return Ivan as result of query for  
 Maria relativeOf ?x  
 when the fact asserted was  
 Ivan childOf Maria

⇒ <P1, owl:inverseOf, P1>



# Physical data representation: RDF vs. RDBMS



Statement		
Subject	Predicate	Object
myo:Person	rdf:type	rdfs:Class
myo:gender	rdfs:type	rdfs:Property
myo:parent	rdfs:range	myo:Person
myo:spouse	rdfs:range	myo:Person
myd:Maria	rdf:type	myo:Person
myd:Maria	rdf:label	"Maria P."
myd:Maria	myo:gender	"F" ○
myd:Maria	rdf:label	"Ivan Jr."
myd:Ivan	myo:gender	"M" ○
myd:Maria	myo:parent	Myd:Ivan
myd:Maria	myo:spouse	myd:John
...		



# Semantic Repositories Gain Users

---

- **3 out of the Top-5 US Defense Contractors**
- **5 out of the Top-10 Pharmaceutical Companies**
- **Dr. Dobbs, includes semantic repositories as one of the major NoSQL trends**
  - May 21, <http://www.ddj.com/cpp/224900500>
- **A Top-5 TV Production Company**
  - Going to use and RDF database as back-end for its website during the Soccer World Championship this year
  - Estimated load of 2 million visitors/day



# SemData Progress

---

- **The Initiative started shaping up in January 2010**
- **SemData@Sofia, March 2010**
  - A round table aimed to shape the SemData initiative
  - 30 participants gathered for 2 days in the poorest EC country
  - 6 (out top-13) of the semantic repository developers
- **SemData@ESWC, May 2010**
  - 4 out of the 13 developers are around the table
  - In few months we 9 out of 13 semantic repository developers
- **SemData@VLDB, Sep 2010**
  - Workshop at the most significant database conference
  - **The call for papers is still open!**

