

Querying Distributed and Diverse Data

SPARQL

Introduction

- Where to start?
 - <http://esw.w3.org/topic/SPARQL>
 - <http://www.w3.org/2007/12/sparql-pressrelease>

Fancy Quotations

- „SPARQL is to the Semantic Web (and, really, the Web in general) what SQL is to relational databases.“
- „SPARQL makes it possible to query information from databases and other diverse sources in the wild, across the Web.“
 - Tim Berners-Lee

SPARQL as W3C standard

- W3C Recommendation
 - since 15th January 2008
- SPARQL is a set of standards:
 - SPARQL Query Language for RDF
 - <http://www.w3.org/TR/rdf-sparql-query/>
 - SPARQL Protocol for RDF
 - <http://www.w3.org/TR/rdf-sparql-protocol/>
 - SPARQL Query Results XML Format
 - <http://www.w3.org/TR/rdf-sparql-XMLres/>

Introduction

- SPARQL facilitates
 - **federation of data**
 - from multiple Web sites (mashups)
 - from multiple enterprise databases (e.g. manufacturing and customer orders and shipping systems)
 - between internal and external systems (e.g. for outsourcing, public Web databases, supply-chain partners)

SPARQL

- SPARQL reference card
 - <http://www.dajobe.org/2005/04-sparql/SPARQLreference-1.7.pdf>
- 14 implementations
 - <http://www.w3.org/2001/sw/DataAccess/impl-report-ql.html>
 - Open RDF Sesame
 - ARQ (Jena)
 - ARC
 - RDF API for PHP
 - Pellet
 - etc.

Basic Facts (1)

- consists of the syntax and semantics for **asking** and **answering queries** against RDF data
- contains capabilities for querying by:
 - **triple patterns,**
 - **conjunctions,**
 - **disjunctions,**
 - **optional patterns**
- supports **constraining queries** by **source RDF graph** and **extensible value testing**
- results of SPARQL queries can be **ordered, limited** and **offset** in number, and **presented in several different forms**

Basic Facts (2)

- several query forms:
 - the **SELECT query** form returns
 - variable bindings
 - the **CONSTRUCT query** form returns
 - an RDF graph
 - the graph is built based on a template which is used to generate RDF triples based on the results of matching the graph pattern of the query

Living endpoints

- <http://esw.w3.org/topic/SpARQLEndpoints>
- URLs of SPARQL exposed over HTTP binding
- dbpedia
 - <http://dbpedia.org/sparql>
- Revyu
 - <http://revyu.com/sparql>
- DBtune
 - <http://purl.org/dbtune/sparql>
- World Factbook
 - <http://www4.wiwiss.fu-berlin.de/factbook/sparql>

Universal Web service

- SPARQL endpoint is a „pure“ web service
- What most of today's web services do?
 - Provide structural data
 - But they aren't efficient at that
 - Expose only limited set of operations
 - Thus more code is required to join the the results
 - SPARQL endpoint can do that all for you

Extensions

- ARQ - Basic Federated SPARQL Query
 - <http://jena.sourceforge.net/ARQ/service.html>

Networked Graphs

- Simon Schenk (Koblenz-Landau University)