INFO/CS 4302
Web Information Systems

FT 2012
Week 9 : Linked Data Technologies (RDF/S, OWL)

- Bernhard Haslhofer -
Plan for today...

• Linked Data Technologies Overview

• RDF

• RDFS, OWL

• Groupwork: movie data in RDF

• Questions, Housekeeping, ...
LINKED DATA TECHNOLOGIES
OVERVIEW
What is Linked Data?

• A method to build a Web of Data
• Architectural style, set of standards
URI

- **Name** and **identify** things (resources)
- **Dereferencable** HTTP URIs

- http://dbpedia.org/resource/The_Shining_(film)
- http://rdf.freebase.com/ns/m/04fjzv
RDF

- A data model for representing data on the Web
- Several statements (triples) form a graph

```
http://dbpedia.org/ontology/Film
  ^ rdf:type
  |   http://dbpedia.org/resource/The_Shining_(film)
      ^ rdfs:label
      | rdfs:label
      v dbpprop:starring
http://xmlns.com/foaf/0.1/Person
  ^ rdf:type
  |   http://dbpedia.org/resource/Jack_Nicholson
      ^ dbpedia-owl:birthDate
      | foaf:name
  v 1937-04-22
```

The Shining (film)

Jack Nicholson
RDF/XML, N3, Turtle, etc.

- **Data formats** for RDF resource representations
- **Used to transfer RDF data between apps**
RDFS

- A **language** for describing the syntax and semantics of **vocabularies** in a machine-understandable way
OWL

• A more expressive (formal) language for defining the syntax and semantics of vocabularies

• Solves RDFS shortcomings but introduces quite some complexity
SKOS

• A language for describing controlled vocabularies (taxonomies, thesauri, classification schemes)
SPARQL

• A query language and protocol for accessing RDF data on the Web

SELECT DISTINCT ?x
WHERE {
    ?x dcterms:subject
}
## Database Systems Analogy...

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Relational Database Management Systems (RDBMS)</th>
<th>Linked Data Technologies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Query</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schema Definition Language</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Representation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identifiers</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Schema Definition Language

- **Purpose**: Describe the structure of data in a database.
- **RDBMS**: Uses SQL or similar languages.
- **Linked Data Technologies**: Uses XML or JSON.

#### Data Representation

- **Purpose**: Store data in a way that can be queried and accessed.
- **RDBMS**: Tables and relations.
- **Linked Data Technologies**: Resource Description Framework (RDF) or JSON-LD.

#### Identifiers

- **Purpose**: Unique identifiers for data entities.
- **RDBMS**: Primary keys.
- **Linked Data Technologies**: URIs or unique identifiers.
# Database Systems Analogy...

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Relational Database Management Systems (RDBMS)</th>
<th>Linked Data Technologies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Query</td>
<td>SQL</td>
<td>SPARQL</td>
</tr>
<tr>
<td>Schema Definition Language</td>
<td>SQL DDL</td>
<td>RDFS / OWL</td>
</tr>
<tr>
<td>Data Representation</td>
<td>Relational Model / Tables</td>
<td>RDF / Graph</td>
</tr>
<tr>
<td>Identifiers</td>
<td>Primary Keys (numeric sequences)</td>
<td>URI</td>
</tr>
</tbody>
</table>
RDF
RDF

• Resource Description Framework

• A graph-based data model to represent data on the Web

• Machine-readability

• Uses URIs to name and identify things
RDF Statements / Triples

- The basic structural element of RDF is the statement / triple

```
http://dbpedia.org/resource/The_Shining_(film)
```

**Subject**
```
http://dbpedia.org/resource/The_Shining_(film)
```

**Property**
```
dbpprop:runtime
```

**Value**
```
146.0
```

```
http://dbpedia.org/resource/The_Shining_(film)
```

**Subject**
```
http://dbpedia.org/resource/The_Shining_(film)
```

**Predicate**
```
rdf:type
```

**Object**
```
dbpedia-owl:Film
```

RDF Statements / Triples

• RDF triples can be merged into a set of triples, constituting a directed graph structure

http://dbpedia.org/resource/The_Shining_(film)

`dbprop:runtime` 146.0

`rdf:type` `dbpedia-owl:Film`
URIs in RDF

• The labels of graph nodes or edges in these slides are either full URIs or use prefixes

• Example:
  – `dbprop:runtime` is a shorthand using the prefix `dbprop`
  – `dbprop` stands for `http://dbpedia.org/property`
  – `dbprop:runtime` therefore qualifies to `http://dbpedia.org/property/runtime`

• Both are equivalent to each other
RDF Serialization

• RDF can be serialized using various syntax formats:
  – NTriples
  – N3/Turtle
  – RDF/XML
  – JSON-LD
  – ....

• The following examples convey the same information
Language Tags

• Literals may carry language tags (@en, @zh)

http://dbpedia.org/resource/The_Shining_(film)

rdfs:label  rdfs:label

闪灵 (电影) @zh
The Shining (film) @en
Typed Literals

- Literals can be typed using arbitrary datatypes
  - XML Schema datatypes
  - custom datatypes
Other RDF features

• Blank Nodes → „Non-URI nodes“

• Containers → „Grouping of resources“

• Collections → „Linked Lists“

• Reification → „Statements about statements“
RDFS, OWL
RDF Vocabulary Description Language (RDFS)

• Extends RDF with the possibility to define
  – classes and associated
  – properties

• Allows different applications to agree on common information models (vocabularies)

• RDF Schema is based on RDF → every RDF Schema document is an RDF document
Classes

http://dbpedia.org/resource/The_Shining_(film)

http://www.w3.org/2000/01/rdf-syntax-ns#type

http://dbpedia.org/ontology/Film

http://www.w3.org/2000/01/rdf-schema#subClassOf

http://www.w3.org/2000/01/rdf-schema#Class

http://www.w3.org/1999/02/22-rdf-syntax-ns#type

http://dbpedia.org/ontology/Work

http://www.w3.org/2000/01/rdf-schema#subClassOf

http://dbpedia.org/ontology/Newspaper
Properties

http://www.w3.org/1999/02/22-rdf-syntax-ns#property

http://www.w3.org/1999/02/22-rdf-syntax-ns#type

dbprop:starring

http://www.w3.org/2000/01/rdf-schema#domain

http://www.w3.org/2000/01/rdf-schema#range

http://dbpedia.org/ontology/Person

http://dbpedia.org/ontology/Film
Other RDFS features

- `rdfs:subPropertyOf`: hierarchical properties
- `rdfs:comment`: human-readable comments
- `rdfs:label`: Human-readable names for resources
- `rdfs:seeAlso`
- `rdfs:isDefinedBy`
RDFS Shortcomings

• No distinction between
  – attributes
  – relationships

• No cardinality constraints (min, max)

→ OWL tries to solve RDFS shortcomings
Web Ontology Language (OWL)

- A language designed to represent rich and complex knowledge about things
- Logic-based
  - verify consistency of defined knowledge
  - make implicit knowledge explicit (inference)
  - driven by AI community
- OWL models can be exchanged as RDF documents
**OWL Class**

- **owl:Class**: defines a group of individuals that belong together because of shared properties
OWL Object Properties

- **owl:ObjectProperty**: properties whose value is an individual
OWL Datatype Properties

- **owl:DatatypeProperty**: properties whose value is a literal
Equality of Individuals

• **owl:sameAs**: Two individuals may be stated to be the same
  – used to link data across data sources
  – controversy on the notion of „sameness“
Class Equivalence

- **owl:equivalentClass**: classes may refer to the same set of individuals
  - used to map between schemas/vocabularies
Property Equivalence

• **owl:equivalentProperty**: two properties may be stated to be equivalent
  – used to map between schemas/vocabularies

```
```
Other OWL Features

• We scratched OWL just on the surface

• More details at:
  http://www.w3.org/TR/owl2-primer/

• A good starting point:
  http://protege.stanford.edu/doc/owl/getting-started.html
How to represent data in RDF

• Design an information model expressing
  – the resource types in your dataset ( = RDFS/OWL classes)
  – their attributes (= RDF/OWL properties)
  – the relationships between them ( = properties)

• Assign names (URIs) to model entities, either by
  – reusing existing terms or
  – defining new, proprietary terms

• Create resources, assign names (URIs), describe them with attributes, and connect them via relationships
How to find vocabulary terms?

• GoodRelations: [http://www.heppnetz.de/projects/goodrelations/](http://www.heppnetz.de/projects/goodrelations/)
• schema.org: [http://schema.org](http://schema.org)
GROUPWORK: MOVIE DATA IN RDF
Instructions

• Form groups of 3
• Take one example movie from the HW dataset (e.g., „The Godfather“)
• Discuss how to represent this movie, its attributes and relationships in RDF
USEFUL APIS / TOOLS
RDF APIs

• Java
  – Jena Semantic Web Framework (http://openjena.org/)
  – Sesame RDF API (http://www.openrdf.org/)

• PHP
  – ARC (http://arc.semsol.org/)

• Ruby
  – RDF.rb: Linked Data for Ruby (http://rdf.rubyforge.org/)

• Python
  – RDFLib (http://www.rdflib.net/)

• C
  – Redland RDF Libraries (http://librdf.org/)
Linked Data debugging

using cURL:

```


curl -iH "Accept: text/n3" http://dbpedia.org/resource/The_Shining_(film)
```
Linked Data debugging

using raptor (http://librdf.org/raptor/):

rapper -o rdfxml
http://dbpedia.org/resource/The_Shining_(film)

rapper http://dbpedia.org/resource/The_Shining_(film)
> ~/Desktop/the_shining.nt
QUESTIONS & HOUSEKEEPING