Agenda

1. Linked (Open) Data

2. Semantic Web and HTML
   - RDFa
   - Microformats
   - Google Knowledge Graph

3. OWL Applications
   - OWL DL Application EDF Energy
   - OWL Profile Application BBC World Cup
   - Semantic Technologies in the Pharmaceutical Industry

4. Summary
Agenda

1. Linked (Open) Data

2. Semantic Web and HTML
   - RDFa
   - Microformats
   - Google Knowledge Graph

3. OWL Applications
   - OWL DL Application EDF Energy
   - OWL Profile Application BBC World Cup
   - Semantic Technologies in the Pharmaceutical Industry

4. Summary
Data in the Web

- more and more data is available in the Web for programmatic access
- often specified using Semantic Web Standards, e.g., the following Linking Open Data (LOD) Initiative
  http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData
- using APIs, e.g. via JSON/REST
- Semantic Web technologies simplify the integration of data from different sources
- combination of data leads to deeper insights
Linked Data in the Web 01.05.2007


TU Dresden  Foundations of Semantic Web Technologies  slide 5 of 51
Linked Data in the Web 31.03.2008

Linked Data in the Web 14.07.2009

Linked Data in the Web 22.09.2010

Linked Data in the Web 19.09.2011

Linked Data in the Web 30.08.2014


TU Dresden 
Foundations of Semantic Web Technologies 
slide 10 of 51
Linked Data Principles*

Linked Data consists mainly of a number of principles for publishing data in the Web:

1. Use URIs as names for things
   - documents, people, locations, concepts, etc.
2. Use HTTP URIs so that people can look up those names
3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)
4. Include links to other URIs, so that they can discover more things.

*http://www.w3.org/DesignIssues/LinkedData.html
5 Star Linked (Open) Data

⭐️ Available on the web (whatever format) but with an open licence, to be Open Data

⭐⭐ Available as machine-readable structured data (e.g. excel instead of image scan of a table)

⭐⭐⭐ as (2) plus non-proprietary format (e.g. CSV instead of excel)

⭐⭐⭐⭐ All the above plus, use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff

⭐⭐⭐⭐⭐ All the above, plus: Link your data to other people’s data to provide context

http://www.w3.org/DesignIssues/LinkedData.html
De-Referencing of an IRI

- These IRIs can then be used also in other documents
- For example in the document `<http://ex.org/jones>`:
  `<#denise> fam:child <#edwin>;
  <http://ex.org/smith#carol>`.
- One can then extract the URL `<http://ex.org/smith>` from
  `<http://ex.org/smith#carol>` **and find information about #carol** from there.
Connection between the IRI of a Thing and IRI of a Source

User Agent

http://www.polleres.net/foaf.rdf#me

HTTP GET RDF

Web Server

http://www.polleres.net/foaf.rdf
Connection between the IRI of a Thing and IRI of a Source

http://dbpedia.org/resource/Gordon_Brown

http://dbpedia.org/data/Gordon_Brown

http://dbpedia.org/page/Gordon_Brown

*HTTP Response Code 303: See Other
Representations

- Information resources can have different representations
- A representation is a stream of bytes in a certain format such as HTML, RDF/XML or JPEG
- Example: an invoice is an information resource that might be represented as printable PDF or as RDF document
- A single resource can have many different representations, e.g., in different formats, resolutions or languages
HTTP Content Negotiation

- Content Negotiation (CN, conneg) is the process of selecting the best representation for a query if several representations are available.

```

```

curl – Tool to send requests to a server or receive responses
-H Custom header to pass to server
-I Show document info only
HTTP Content Negotiation

```
$ curl -I -H "Accept: text/html"
   http://dbpedia.org/resource/Gordon_Brown
HTTP/1.1 303 See Other
Date: Mon, 04 Feb 2013 10:37:10 GMT
Content-Type: text/html; charset=UTF-8
Content-Length: 0
Connection: keep-alive
Server: Virtuoso/06.04.3132 (Linux) [...]
Accept-Ranges: bytes
Location: http://dbpedia.org/page/Gordon_Brown
```
HTTP Content Negotiation

HTTP/1.1 303 See Other
Date: Mon, 04 Feb 2013 10:36:59 GMT
Content-Type: application/rdf+xml; qs=0.95
Content-Length: 0
Connection: keep-alive
Server: Virtuoso/06.04.3132 (Linux) [...] 
Accept-Ranges: bytes 
TCN: choice
Vary: negotiate,Accept
Content-Location: /data/Gordon_Brown.xml
Link: <http://mementoarchive.lanl.gov/dbpedia [...] 
Location: http://dbpedia.org/data/Gordon_Brown.xml
Linked Data Applications: Minimal Architecture
Linked Data Summary

Semantic technologies simplify the access to data:

- Facts regarding Berlin?
  - http://de.dbpedia.org/resource/Berlin

- Information about Queen
  - BBC Music: http://www.bbc.co.uk/music/artists/0383dadf-2a4e-4d10-a46a-e9e041da8eb3
  - MusicBrainz: http://musicbrainz.org/artist/0383dadf-2a4e-4d10-a46a-e9e041da8eb3.html

- Data integration gives additional benefits
Linked Data Tools

- **Tabulator Browser PlugIn/Ajax Scripts:**
  http://www.w3.org/2005/ajar/tab

- **Semantic Web Client Library (Querying the complete Semantic Web with SPARQL):**
  http://wifo5-03.informatik.uni-mannheim.de/bizer/ng4j/semwebclient/

- **D2R Server: Accessing databases with SPARQL and as Linked Data:**
  http://d2rq.org/d2r-server

- **Data cleaning & linking to Freebase:**
  https://github.com/OpenRefine (was Google Refine)

- **RDF Export for Google Refine:**
Agenda

1. Linked (Open) Data

2. Semantic Web and HTML
   - RDFa
   - Microformats
   - Google Knowledge Graph

3. OWL Applications
   - OWL DL Application EDF Energy
   - OWL Profile Application BBC World Cup
   - Semantic Technologies in the Pharmaceutical Industry

4. Summary
RDFa Example

- Integration of RDF in (X)HTML documents

All content on this site is licensed under <a href="http://creativecommons.org/licenses/by/3.0/">a Creative Commons License</a>.

**versus**

All content on this site is licensed under <a rel="license" href="http://creativecommons.org/licenses/by/3.0/">
  a Creative Commons License</a>.

license

<http://example.org/a.html>
RDFa Example 2

```html
<div>
  <h2>The trouble with Bob</h2>
  <h3>Alice</h3>
  ...
</div>

versus

```html
<div xmlns:dc="http://purl.org/dc/elements/1.1/">
  <h2 property="dc:title">The trouble with Bob</h2>
  <h3 property="dc:creator">Alice</h3>
  ...
</div>
```
RDFa Example 3

```html
<div>
  <p>Alice Birpemswick</p>
  <p>Email: <a href="mailto:alice@example.com">alice@example.com</a></p>
  <p>Phone: <a href="tel:+1-617-555-7332">+1 617.555.7332</a></p>
</div>
```

versus

```html
<div typeof="foaf:Person"
  xmlns:foaf="http://xmlns.com/foaf/0.1/">
  <p property="foaf:name">Alice Birpemswick</p>
  <p>Email: <a href="mailto:alice@example.com" rel="foaf:mbox">alice@example.com</a></p>
  <p>Phone: <a href="tel:+1-617-555-7332" rel="foaf:phone">+1-617-555-7332</a></p>
</div>
```
Applications of RDFa

- Google filters, for example, RDFa terms and uses them to improve the presentation of search results
Applications of RDFa

- Google filters, for example, RDFa terms and uses them to improve the presentation of search results
Microformats

Microformats are simple and open data formats based on existing standards (XHTML)

- Uses POSH (Plain Old Semantic HTML), i.e., HTML Tags that do not specify the presentation (bold, i), but that have semantics (abbr, acronym, title, ...)
- Use of semantic CSS class names
  - not: <span class="blueText">...</span>
  - but: <span class="submenu">...</span>
- Special vocabularies for the markup in some domains
- Considers the roles and semantics of the elements
hRecipe – Microformat for Recipes

- For the semantic annotation of web pages for recipes
- Allows for
  - searching for recipes with certain ingredients
  - automatic grouping of recipes
  - finding quick recipes (short preparation time)
- Mapping into RDFa exists (hrecipe-rdf)

Example

```html
<div class="hrecipe">
  <h1 class="fn">French Fries</h1>
  <p class="summary">French Fries ...</p>
  <p>
    Contributed by <span class="author">Tom</span> and the <span class="vcard author">
      <a class="url fn" href="...">Cooky Gang</a>
    </span>.
  </p>
...
Example

...<p>Published <span class="published"> <span class="value-title" title="2008-10-14T10:05:37-01:00"/> 14. Oct 2008</span></p> <h2>Ingredients</h2> <ul> <li class="ingredient"> <span class="value">500</span> <span class="type">gr</span> potatoes. </li> ... </ul> ...</div>
Example in the Web

- The Recipe Depository http://www.therecipepository.com/
- SAPO Sabores http://sabores.sapo.pt/
- ITV Food http://www.itv.com/food/
- Epicurious.com http://www.epicurious.com/
- http://foodnetwork.com/
- Plan to Eat http://www.plantoeat.com/recipe_book
- essen & trinken http://www.essen-und-trinken.de/
  hRecipe-conform meta data in RDF
- ...
Micro Data and schema.org

- Joint project of Google, Microsoft and Yahoo!
- Compromise between the extensibility of RDFa and the simplicity of microformats
- Goal: better understanding of the contents of web pages and, as a result, a better presentation of search results
Google Knowledge Graph

- The US version of Google now also uses structured data (from Freebase)
- For the disambiguation of search terms and direct presentation of relevant information
- Considers search terms no longer as simple strings (but as designator for things)
Search Results:

**The University of Louisiana at Monroe:**
www.ulm.edu/  
Degree programs are offered by the seven Colleges and the Graduate School.

**Ulm - Wikipedia, the free encyclopedia**
en.wikipedia.org/wiki/Ulm  
Ulm is a city in the federal German state of Baden-Württemberg, situated on the River Danube. The city, whose population is estimated at 120,000 (2006), forms ...

**ULM Warhawks Athletics**
www.ulmwarhawks.com/  
The official athletics site for the University of Louisiana Monroe Warhawks. The home for all ULM sports schedules, news, rosters and stats: football, men's ...

**English - Stadt Ulm - Start**
www.ulm.de/start.3080.htm  
14 Jul 2005 – City of Ulm online. Jump for: To mainnavigation . to content . to change language . to search . to "much asked for.

**Images for Ulm - Report images**

**Homepage - Universität Ulm**
www.uni-ulm.de/en/  
Die Fakultät für Naturwissenschaften der Universität Ulm hat am Freitag Professor Jürgen Münck, Präsident der Helmholtz-Gemeinschaft und damit der größten ...

**Ulm Photonics**

**University of Louisiana at Monroe**
The University of Louisiana at Monroe is a coeducational public university in Monroe, Louisiana, United States and part of the University of Louisiana System. Wikipedia

**Nickname:** Warhawks
**Address:** 700 University Avenue, Monroe, Louisiana, 71209
**Mascot:** Ace the Warhawk
**Enrollment:** 8,811 (2010)
**Acceptance rate:** 72.9% (2010)
**Colors:** Gold, Maroon

**People also search for**
University of Louisiana at Lafayette  
Louisiana Tech University  
Louisiana State University  
Northwestern State University  
Grambling State University

See results about
Ulm  
Ulm is a city in the federal German state of Baden-Württemberg, situated on the River Danube. ...

University of Ulm  
The Ulm University is a public university in the city of Ulm, in the South German state of ...
University of Louisiana at Monroe

The University of Louisiana at Monroe is a coeducational public university in Monroe, Louisiana, United States and part of the University of Louisiana System. Wikipedia

Nickname: Warhawks
Address: 700 University Avenue, Monroe, Louisiana, 71209
Mascot: Ace the Warhawk
Enrollment: 8,811 (2010)
Acceptance rate: 72.9% (2010)
Colors: Gold, Maroon

People also search for
University of Louisiana at Lafayette
Louisiana Tech University
Louisiana State University
Northwestern State University
Grambling State University

See results about
Ulm
Ulm is a city in the federal German state of Baden-Württemberg, situated on the River Danube. ...

University of Ulm
The Ulm University is a public university in the city of Ulm, in the South German state of ...
Search

5 personal results. 29,600,000 other results.

- Taj Mahal - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Taj_Mahal
The Taj Mahal is a white marble mausoleum located in Agra, India. It was built by Mughal emperor Shah Jahan in memory of his third wife, Mumtaz Mahal.

- Taj Mahal (musician) - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Taj_Mahal_(musician)
Henry Saint Clair Fredericks (born May 17, 1942), who uses the stage name Taj Mahal, is an American Grammy Award winning blues musician. He often ...

- Atlantic City New Jersey Casino Hotels | Trump Taj Mahal, Atlantic City
www.trumpatlas.com
For five star Atlantic City casino hotels, Trump Taj Mahal Casino Hotel offers restaurants, nightlife and a casino floor you can't resist. Explore our Atlantic City ...

- Images for Taj Mahal - Report Images

Welcome To Official Website of Taj Mahal-U.P. Tourism
www.tajmahal.gov.in
Taj Mahal at Agra, India is one of the wonders of the world. Find information related to Taj Mahal travel and tourism in this article.
Score: 27 / 30 - 326 Google reviews - Write a review

- Taj Mahal
www.tajmahal.org.uk
Taj Mahal at Agra, India is one of the wonders of the world. Find information related to Taj Mahal travel and tourism in this article.

- Taj Mahal - Welcome to tajblues.com HOME
www.tajblues.com

People also search for

Taj Mahal
Musician
Henry Saint Clair Fredericks, who uses the stage name Taj Mahal, is an American Grammy Award ...
Agenda

1. Linked (Open) Data

2. Semantic Web and HTML
   - RDFa
   - Microformats
   - Google Knowledge Graph

3. OWL Applications
   - OWL DL Application EDF Energy
   - OWL Profile Application BBC World Cup
   - Semantic Technologies in the Pharmaceutical Industry

4. Summary
Application Areas of OWL

- OWL DL mainly is used outside the Semantic Web
- Many applications in medicine and life sciences
- Terminologies are traditionally popular there
- Keyword indexing of documents
- Semantic annotations of research data (e.g., gene sequences)
- Classification used in health records and for statistics
Example Ontologies in OWL

- OBO Foundry: The Open Biological and Biomedical Ontologies
- BioPortal ontologies
  - Terms for the electronic patient record
  - Annotation of gene sequences
  - Research into new drugs
- GO Gene Ontology
- ICD International Classification of Diseases
- FMA Formal Model of Anatomy
- ...
Use of OWL in the EDF Energy Management Advisor

Weather
Consumption
Building
... 50 parameters

EMA

Tips
EMA Ontology

- Ontology models the domain and situation of customers
- First modelling was strictly following existing binary decision diagrams (simplified)
- Reasoner “recognises” the situation of a customer
- Certain situations correspond to tips
- Original ontology used nominals and role chains
- Was difficult to comprehend and reasoner performance was not optimal
EMA Ontology Improvements

- Customers are directly modelled in an ABox
- Per customer only simple ABox facts are loaded
- Constructors that are problematic for reasoning are avoided: nominals, role chains
- TBox modelling was simplified
- Ontology now allows incremental reasoning
- Customers can be classified independent of each other in different reasoner instances
- Used for about 30,000 customers in France
BBC Website for the Football World Cup 2010

- Ontology describes how facts about the world cup relate to each other
- Such meta data are saved as RDF triples
- For example, “Frank Lampard” is part of “England Squad” or “England Squad” competed in “Group C” of the “FIFA World Cup 2010”
“The underlying publishing framework does not author content directly; rather it publishes data about the content - metadata. The published metadata describes the world cup content at a fairly low-level of granularity, providing rich content relationships and semantic navigation. By querying this published metadata we are able to create dynamic page aggregations for teams, groups and players.”

Jem Rayfield, Senior Technical Architect, BBC News and Knowledge
BBC Website for the Football World Cup 2010

- OWL inference used to enrich the data (forward chaining) and SPARQL used for queries
- Ontology contains texts contributed by journalists: stories, blogs, profiles, pictures, videos and statistics
- Journalistic contributions are automatically classified (NLP techniques) and manually tagged
- Statistics and game results from other sources are imported from XML and mapped to ontological concepts
- Web pages are automatically assembled and contain relevant links
- Approach also used for Olympia 2012
Semantic Web Technologies Meet Pharmaceutical Data

Phil Ashworth presents at the 2. European Semantic Technology Conference: http://videolectures.net/estc08_ashworth_swtpdi/
Agenda

1. Linked (Open) Data
2. Semantic Web and HTML
   - RDFa
   - Microformats
   - Google Knowledge Graph
3. OWL Applications
   - OWL DL Application EDF Energy
   - OWL Profile Application BBC World Cup
   - Semantic Technologies in the Pharmaceutical Industry
4. Summary
Summary

- The amount of available machine processable data grows continuously
- Semantics is needed to integrate data from different sources
- Query and visualisation of data provides added value

- The processing and querying data from different sources increases the transparency and facilitate research (tests of hypotheses becomes easier)