COMP318

Ontologies and Semantic Web

RDF - Part 5

Dr Valentina Tamma

V.Tamma@liverpool.ac.uk

Where were we

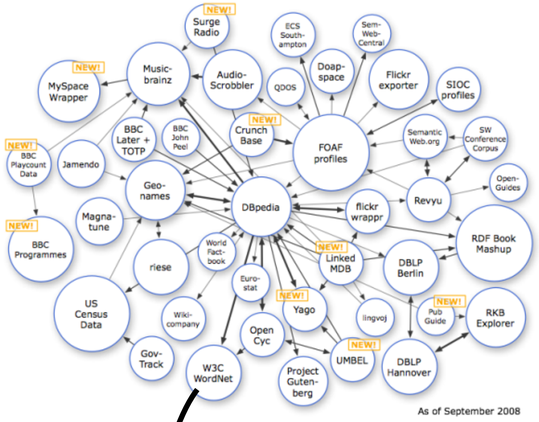
- RDF:
  - data model for sharing information on the Web
  - triples
  - RDF/XML and Turtle

RDFa

- One of the use cases for RDF is to mark up the content of HTML pages
  - The RDFa syntax was introduced to meet this requirement
  - RDFa embeds RDF within the attributes of HTML tags

RDFa

- RDFa is a W3C recommendation that adds a set of attribute level extensions to X(HTML)/XML for embedding rich metadata.
  - it supports the notions of namespaces and URIs
    - it allows the mixing of vocabularies as in RDF
  - it offers a flexible framework for using Resources of type URI or Literal
  - it is a complete serialisation of RDF
    - `http://www.w3.org/TR/rdfa-syntax/`



Picture from [4]

Web of Data

Semantic Web

Semantic Web

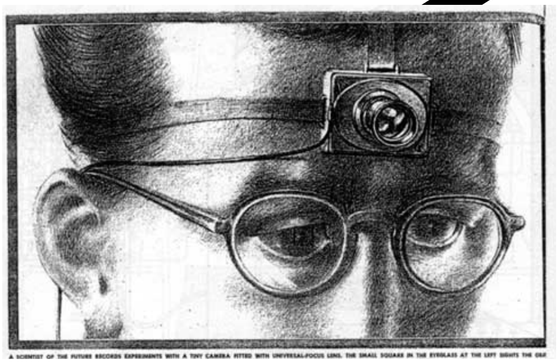
Semantic Annotations

Web

Hypermedia

Hypertext

?



Picture from [3]

“As We May Think”, 1945

- [3] V. Bush "As We May Think" The Atlantic Monthly, July, 1945. Available online: http://www.theatlantic.com/doc/194507/bush
- [4] Linked Data, http://linkeddata.org/

RDFa: embedding in HTML

- RDFa (RDF attribute) allows the embedding of semantic information in existing (X)HTML documents
  - extends (X)HTML a bit by:
    - defining general attributes to add metadata to any elements
    - provides an almost complete “serialisation” of RDF in XHTMLdesigned to enrich existing pages that have already limited semantics based on hyperlinks and tag layout
    - `div` and `span`
  - the RDF data is embedded within the HTML DOM
    - existing content can be annotated with RDFa by simply modifying the HTML document

Microformats

- An approach to add meaning to HTML elements and to make data structures in HTML pages explicit.

- "*Designed for **humans first** and **machines second**, microformats are a set of simple, open data formats built upon existing and widely adopted standards. Instead of throwing away what works today, microformats intend to solve simpler problems first by adapting to current behaviours and usage patterns (e.g. XHTML, blogging).*"

Microformats

- Are highly correlated with semantic (X)HTML / “Real world semantics” / “Lowercase Semantic Web”
- Real world semantics (or the Lowercase Semantic Web) is based on three notions:
  - Adding of simple semantics with microformats (small pieces)
  - Adding semantics to the today’s Web instead of creating a new one (evolutionary not revolutionary)
  - Design for humans first and machines second (user centric design)
- A way to combine human with machine-readable information.
- Provide means to embed structured data in HTML pages.
- Build upon existing standards.

Microformats

- Solve a single, specific problem (e.g. representation of geographical information, calendaring information, etc.).
  - Provide an “API” for your website.
- Build on existing (X)HTML and reuse existing elements.
  - Work in current browsers.
- Follow the DRY principle (“Don’t Repeat Yourself”).
- Compatible with the idea of the Web as a single information space.

Main principles of RDFa

RDFa is a serialization of RDF embedded in XHTML, HTML, or XML in general

- Most of the data on the web are in (X)HTML:
  - new content generated every day
  - how do we get structured data from that info?

- Especially when authors of the “traditional web” don’t like to generate RDF/XML files separately
  - RDF/XML is complex
  - it requires a separate storage, generation, etc. mechanism
    - that is also valid for, e.g., Turtle
    - but even when authoring with a text editor, creating an extra file is a load

What does this mean in practice?

- The same (X)HTML file:
  - is used, unchanged, by browsers
    - they ignore attributes they do not know
  - can be used by specialised processors (or APIs) to extract RDF triples

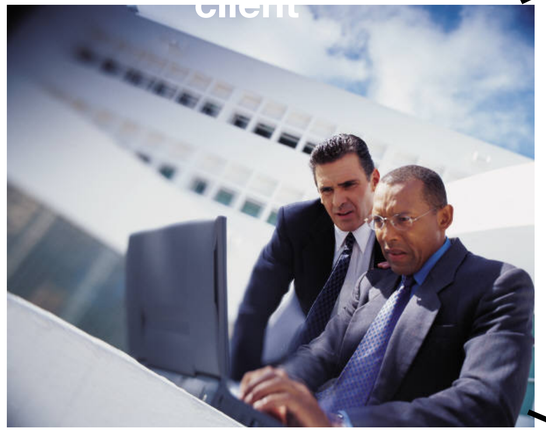
Baron Way apartment for sale

- Pure HTML page
  - No machine readable description

```
<html>
   <body>
    <H1> Baron Way Apartment for Sale</H1>
       The Baron Way Apartment has three bedrooms and is located in
       the family friendly Baron Way Building. The Apartment is located
       in the north of Amsterdam.
   </body>
</html>
```

Typical usage pattern

Request for
`http://www.w3.org/ns/entailment/data/RDFS`



Request for
`http://www.w3.org/ns/entailment/data/RDFS.ttl`

http://www.w3.org/ns/entailment/data/RDFS.html

W3C Semantic Web

Unique identifier for *RDFS Entailment*.

"http://www.w3.org/ns/entailment/RDFS" is the URI. The specification for the RDFS entailment is part of the RDF Semantics W3C Recommendation.

For more information about RDF, please refer to the the RDF Concepts and Abstract Syntax Recommendation .

Ivan Herman, ivan@w3.org, W3C, Semantic Web Activity Lead, 2009-05-03

http://www.w3.org/ns/entailment/data/RDFS.ttl

```
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .


ent:RDFS a ent:Entailment ;
     dc:creator <http://www.ivan-herman.net/foaf#me> ;
     dc:date "2009-05-03" ;
     dc:description "Unique identifier for RDFS Entailment" ;
     rdfs:comment "The specification for the RDFS entailment is part of the RDF Semantics
W3C Recommendation." ;
     rdfs:isDefinedBy <http://www.w3.org/TR/2004/REC-rdf-mt-20040210/#rdfs_entailment> ;
     rdfs:seeAlso <http://www.w3.org/TR/2004/REC-rdf-mt-20040210/> .

<http://www.w3.org/ns/entailment/data/RDFS.html> dc:title "Information Resource RDFS
Entailment" ;
     xhv:stylesheet <http://www.w3.org/StyleSheets/TR/base> .
```

Where does the Turtle content come from?

- The triples are embedded in the HTML file
  - a client may know how to extract RDF triples directly from that file; or
  - an online “distiller” service is used; or
  - the server is set up to generate the Turtle file automatically

- However... the content is created only once!

COMP318

Ontologies and Semantic Web

End of RDF - Part 5

Dr Valentina Tamma

V.Tamma@liverpool.ac.uk