

COMP318: RDF serialisations

`www.csc.liv.ac.uk/~valli/Comp318`



Dr Valentina Tamma

Room: Ashton 2.12

Dept of computer science

University of Liverpool

`V.Tamma@liverpool.ac.uk`

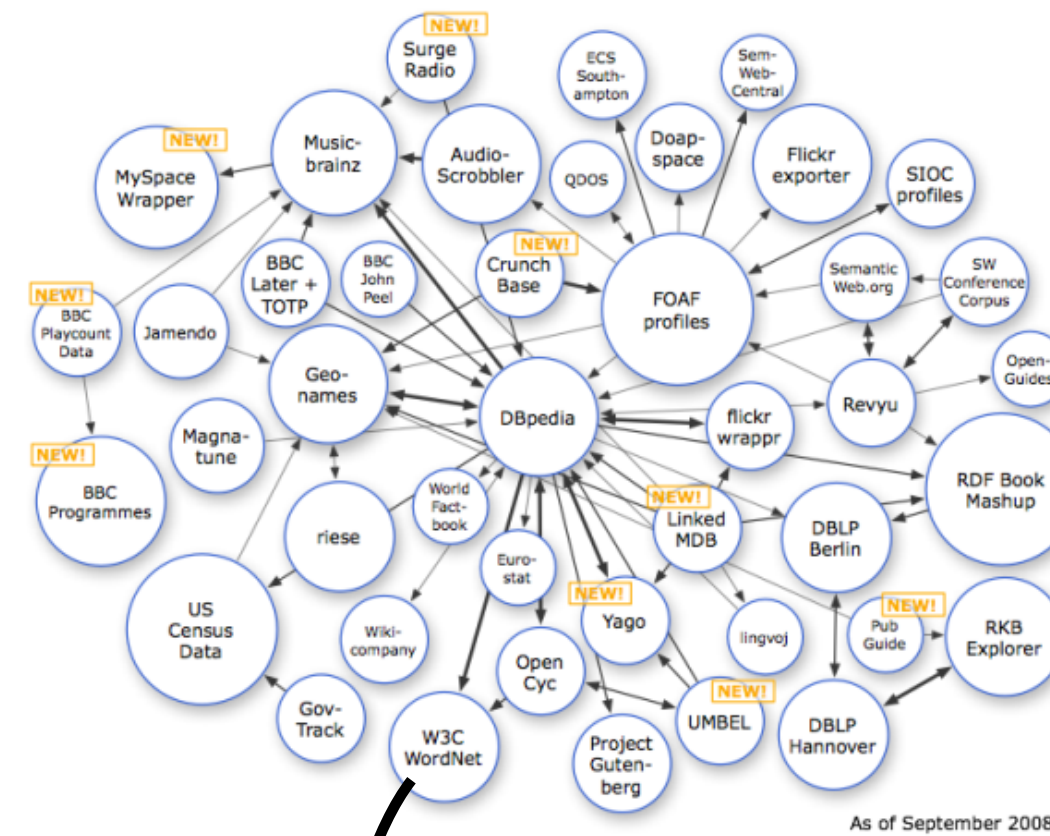
Where were we

- Serialization:
 - RDF document published in a chosen syntax
 - XML/RDF, Turtle and N-triples, **RDFa**, RDF/JSON
- Introduction to RDFa

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



Picture from [4]

?

Web of Data

Semantic Web

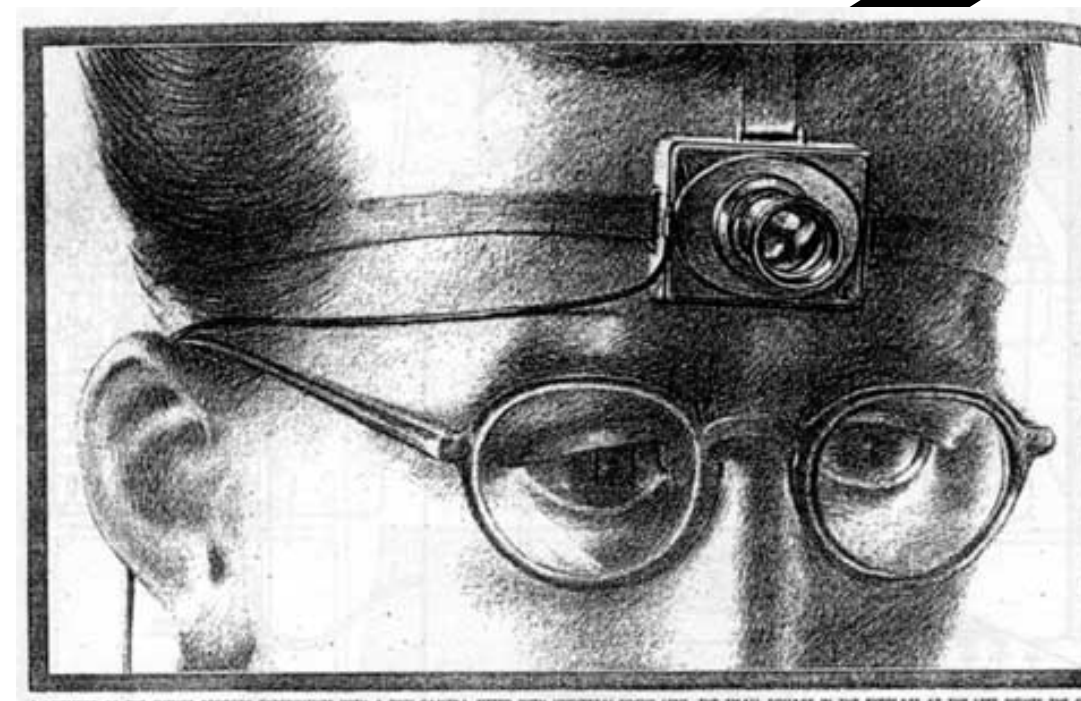


Web

Hypermedia

Semantic Annotations

Hypertext



“As We May Think”, 1945

Picture from [3]

- [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/>

Microformats

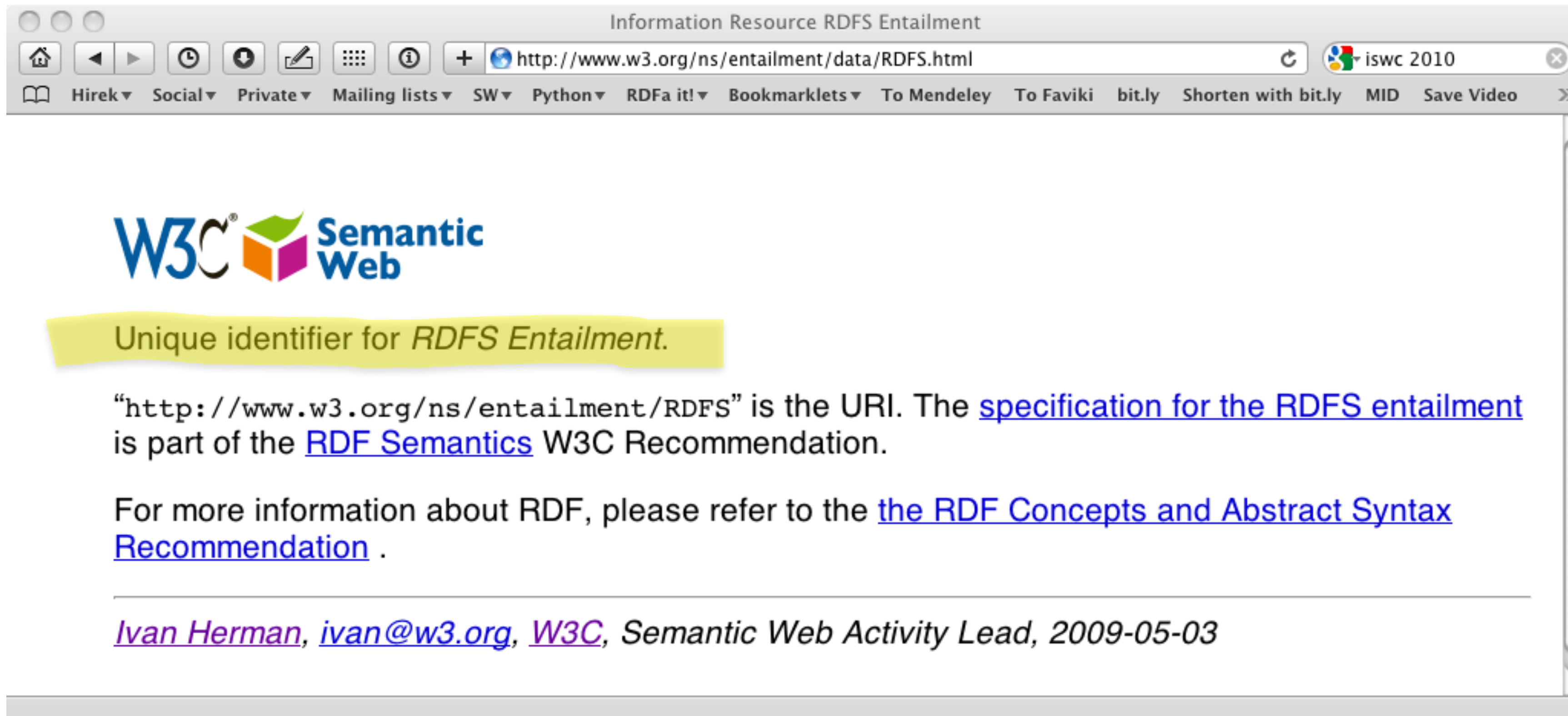
- An approach to add meaning to HTML elements and to make data structures in HTML pages explicit.
- “Designed for **humans first** and **machines second**, microformats are a set of simple, open data formats built upon existing and widely adopted standards. Instead of throwing away what works today, 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.



The source and generated RDF...

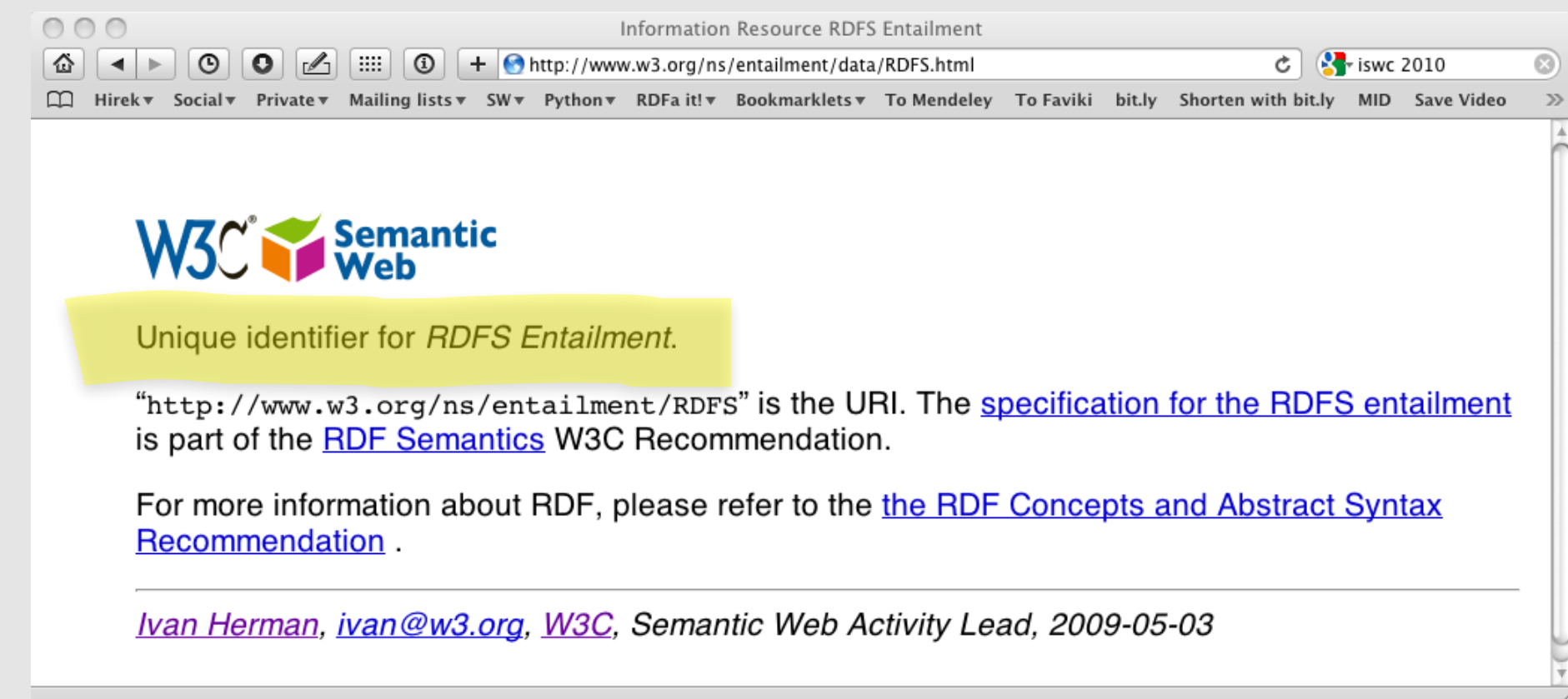
```
<p about="http://www.w3.org/ns/entailment/RDFS"
```

```
property="http://purl.org/dc/terms/description">
```

Unique identifier for

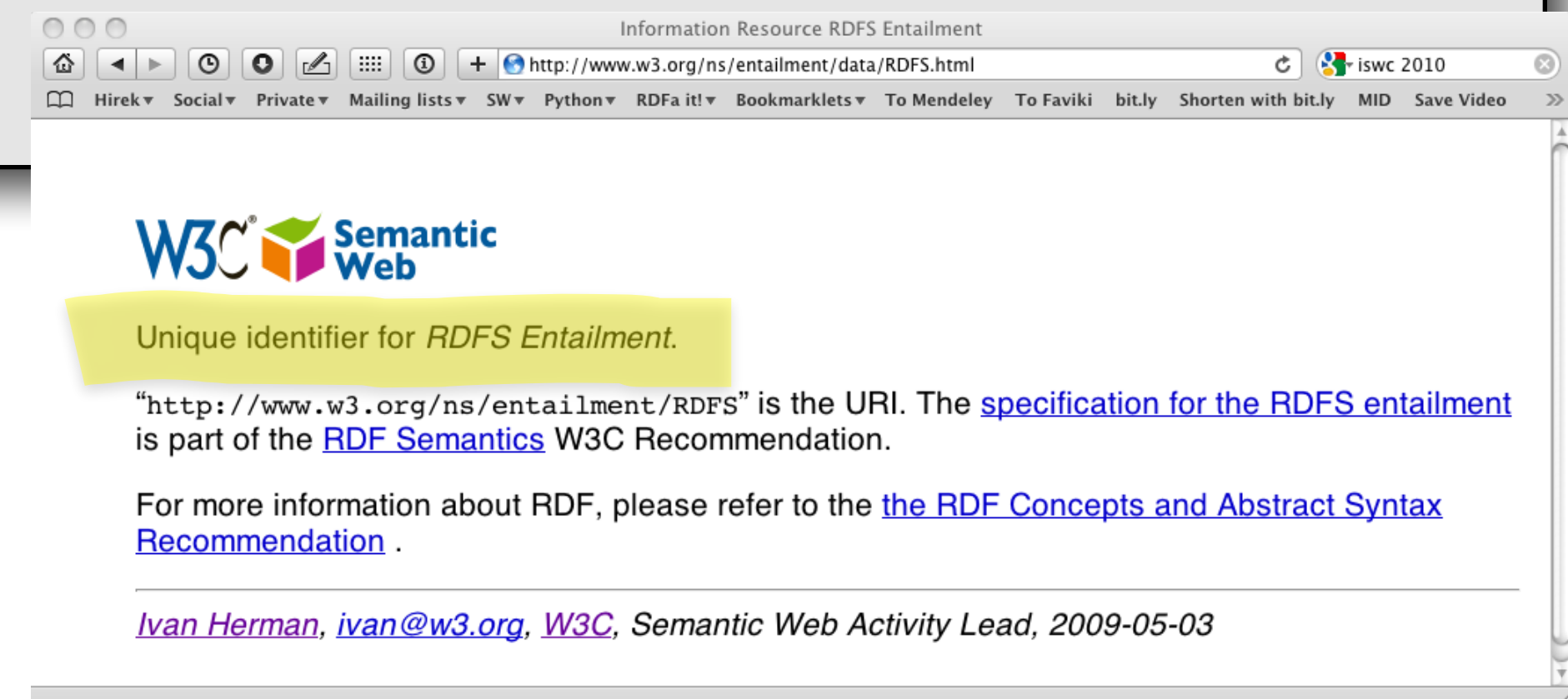
RDFS Entailment.

```
</p>
```



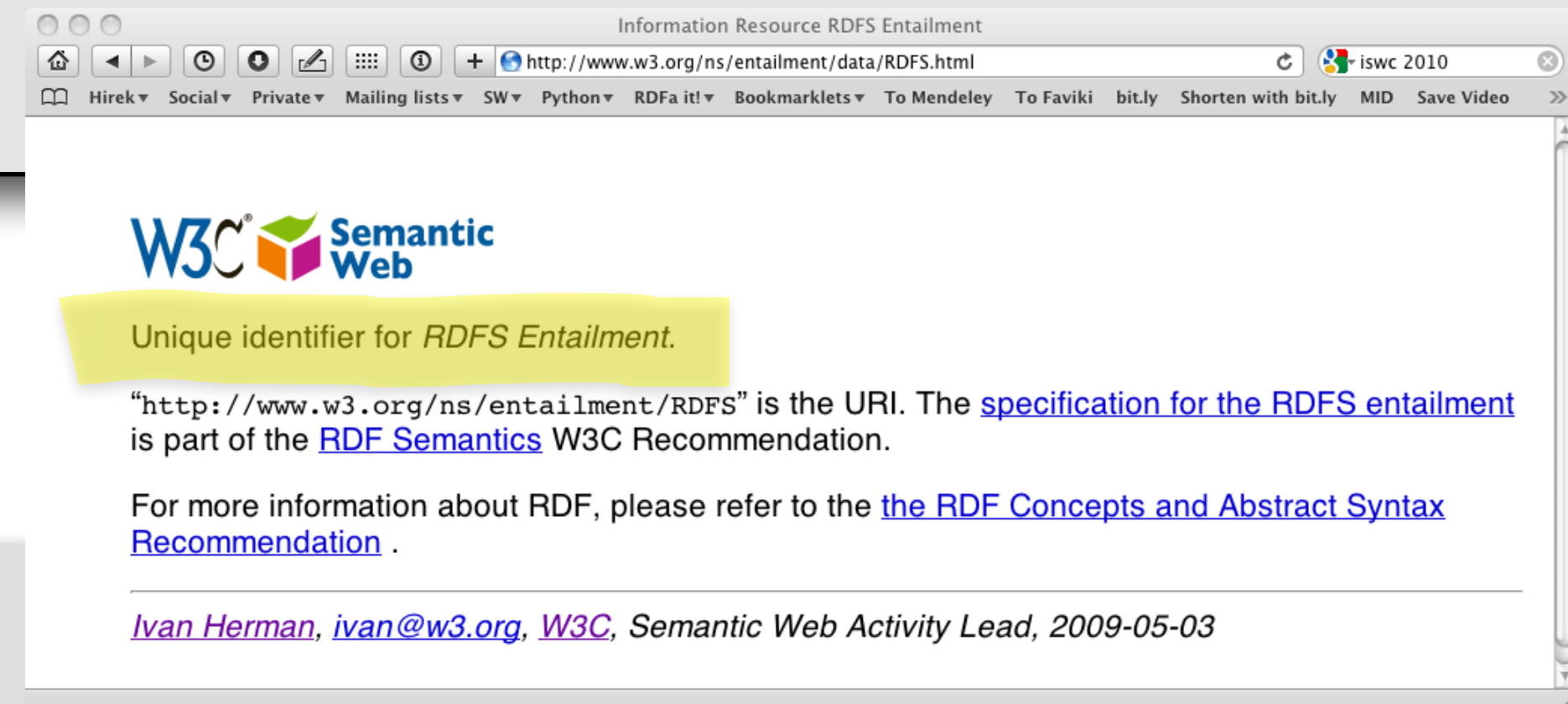
The source and generated RDF...

```
<p about="http://www.w3.org/ns/entailment/RDFS"  
  property="http://purl.org/dc/terms/description">  
  Unique identifier for  
  <em>RDFS Entailment</em>.  
</p>
```



The source and generated RDF...

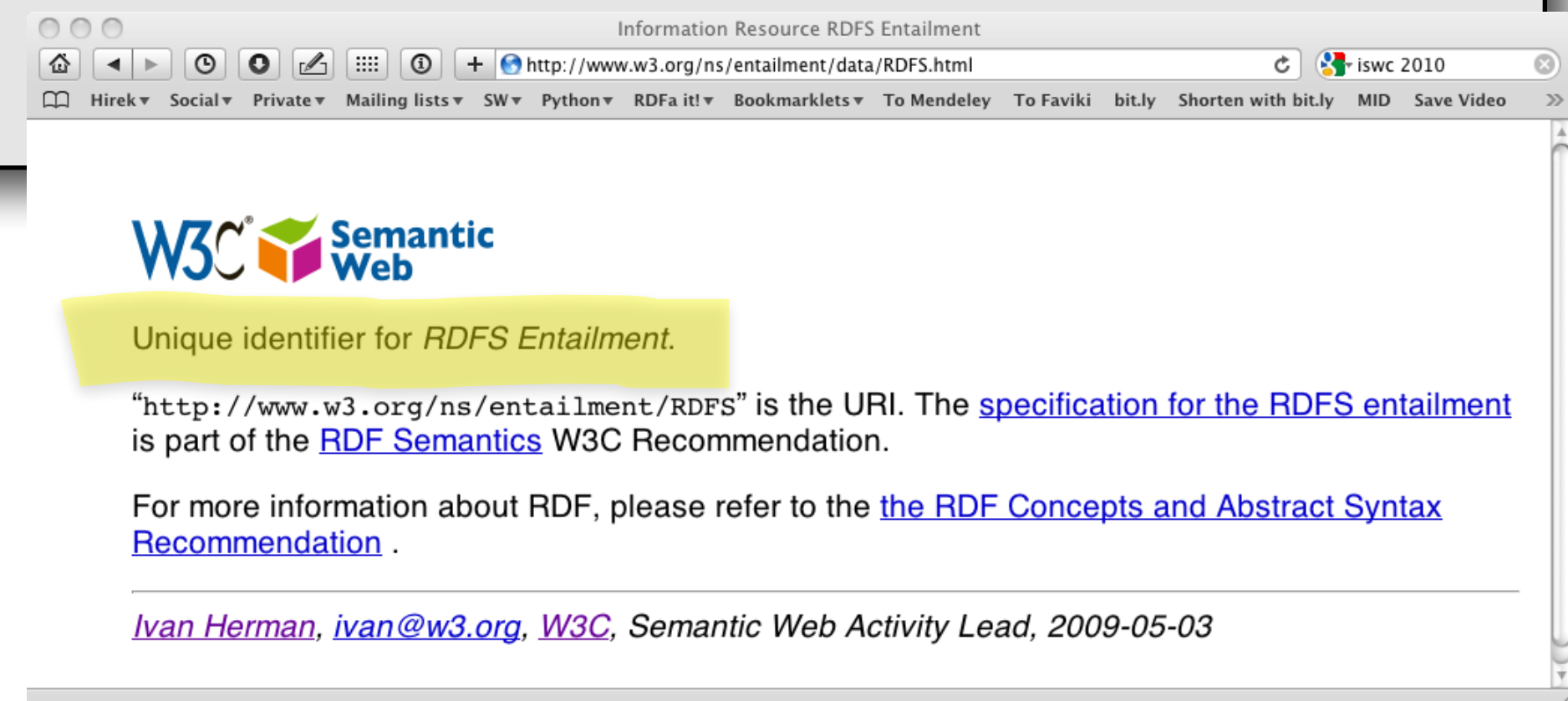
```
<p about="http://www.w3.org/ns/entailment/RDFS"  
  property="http://purl.org/dc/terms/description">  
  Unique identifier for  
  <em>RDFS Entailment</em>.  
</p>
```



```
<http://www.w3.org/ns/entailment/RDFS>
```

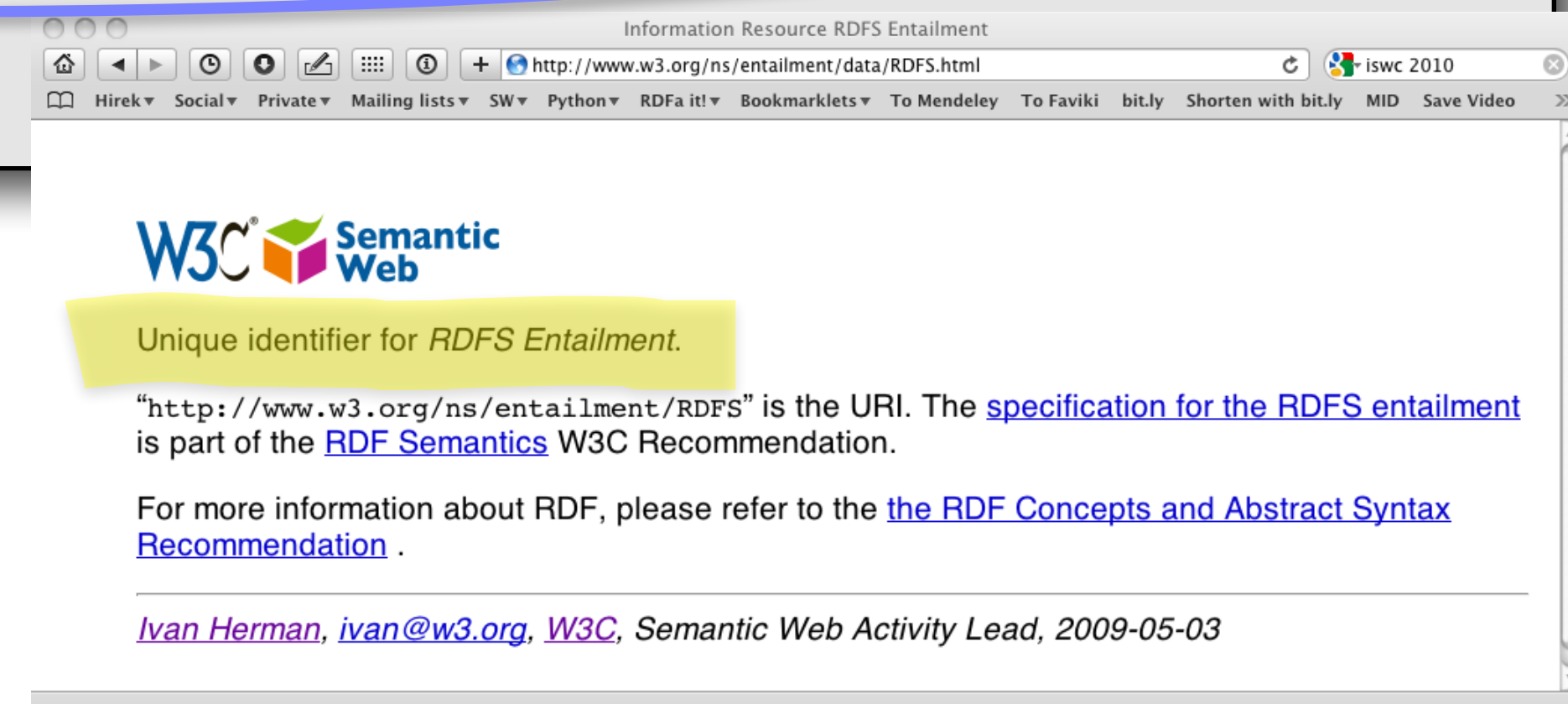
The source and generated RDF...

```
<p about="http://www.w3.org/ns/entailment/RDFS"  
  property="http://purl.org/dc/terms/description">  
  Unique identifier for  
  <em>RDFS Entailment</em>.  
</p>
```



The source and generated RDF...

```
<p about="http://www.w3.org/ns/entailment/RDFS"  
  property="http://purl.org/dc/terms/description">  
  Unique identifier for  
  RDFS Entailment.  
</p>
```



The source and generated RDF...

```
<p about="http://www.w3.org/ns/entailment/RDFS"  
  property="http://purl.org/dc/terms/description">  
  Unique identifier for  
  <em>RDFS Entailment</em>.  
</p>
```

```
<http://www.w3.org/ns/entailment/RDFS>  
<http://purl.org/dc/terms/description>  
"Unique identifier for RDFS Entailment ."
```


The source and generated RDF...

```
<a about="http://www.w3.org/ns/entailment/RDFS"  
    rel="http://www.w3.org/2000/01/rdf-schema#seeAlso">  
    href="http://www.w3.org/TR/2004/REC-rdf-mt-20040210/">  
  
</a>
```

The source and generated RDF...

```
<a about="http://www.w3.org/ns/entailment/RDFS"  
rel="http://www.w3.org/2000/01/rdf-schema#seeAlso">  
href="http://www.w3.org/TR/2004/REC-rdf-mt-20040210/">  
</a>
```

```
<http://www.w3.org/ns/entailment/RDFS>  
  <http://www.w3.org/2000/01/rdf-schema#seeAlso>
```

The source and generated RDF...

```
<a about="http://www.w3.org/ns/entailment/RDFS"  
rel="http://www.w3.org/2000/01/rdf-schema#seeAlso">  
href="http://www.w3.org/TR/2004/REC-rdf-mt-20040210/">  
</a>
```

```
<http://www.w3.org/ns/entailment/RDFS>  
<http://www.w3.org/2000/01/rdf-schema#seeAlso>  
<http://www.w3.org/TR/2004/REC-rdf-mt-20040210/>
```

Just compare

```
<http://www.w3.org/ns/entailment/RDFS>  
  <http://purl.org/dc/terms/description>  
    "Unique identifier for RDFS Entailment ."  
<http://www.w3.org/ns/entailment/RDFS>  
  <http://www.w3.org/2000/01/rdf-schema#seeAlso>  
    <http://www.w3.org/TR/2004/REC-rdf-mt-20040210/> .
```

```
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .  
@prefix dcterms: <http://purl.org/dc/terms/> .  
  
<http://www.w3.org/ns/entailment/RDFS>  
  rdfs:seeAlso <http://www.w3.org/TR/2004/REC-rdf-mt-20040210/> ;  
  dcterms: description "Unique identifier for RDFS Entailment ."
```

The Turtle in RDFa

- Is that it?
 - The combination of **@about** with **@rel** / **@property** and possibly **@href** covers most of what we need...
 - but this is too complex for authors
- Go Turtle:
 - Use compact URIs when possible
 - Make use of the natural structure for
 - shared subjects
 - shared predicates
 - create blank nodes
 - ...

Compact URIs = CURIE's

- Just like in Turtle
 - define a prefix via `@prefix`
 - use `prefix:reference` to abbreviate a URI

Compact URIs = CURIE's

```
<html>
...
<p about="http://www.w3.org/ns/entailment/RDFS"
  property="http://purl.org/dc/terms/description">
  Unique identifier for
  <em>RDFS Entailment</em>.
</p></html>
```

```
<html prefix = "dcterms:http://purl.org/dc/terms/">
...
<p about="http://www.w3.org/ns/entailment/RDFS"
  property="http://purl.org/dc/terms/description">
  Unique identifier for
  <em>RDFS Entailment</em>.
</p></html>
```

RDFa supported attributes

- **xmlns**: a prefix and qualified URL defining a namespace for a document;
- **rel**: a white-space separated list of reserved keywords or CURIes (**C**ompact **U**RI**s**) that details predicates between resources
 - No literals!
 - CURIes can be considered a datatype found both in XML and non-XML grammars
 - syntax: [isbn:0393315703]
 - the [] prevent ambiguity between CURIes and regular URIs.
 - QNames can be considered a type of CURIes
- **rev**: similar to rel, but traverses the predicate in the opposite direction wrt **rel**

RDFa supported attributes

- **about**: a resource URI or CURIE used to represent the subject in an RDF triple
- **property**: a white-space separated list of CURIEs representing predicates between a subject and a plain literal.
- **rel**: represents predicates between a subject and another resource
 - ... and many more
- <http://www.w3.org/TR/rdfa-syntax/>
- <http://www.w3.org/TR/xhtml1-rdfa-primer/>

Consuming RDFa

- Various search engines begin to consume RDFa
 - Google, Yahoo, ...
 - they may specify which vocabularies they “understand”
 - this is still an evolving area
- Facebook’s “social graph” is based on RDFa

Google's rich snippet

- Embedded metadata (microformat or RDFa) is used to improve search result page
 - at the moment only a few vocabularies are recognised, but that will evolve over the years



Who uses it

- A number of popular sites publish RDFa as part of their normal pages:
 - Google.com
 - Youtube.com
 - Facebook.com
 - Wikipedia.org
 - Yahoo.com
 - Amazon.com
 - Reddit.com
 - Netflix.com
 - Creative Commons snippets are in RDFa