

CMSC 691s

RDFa and GRDDL

- Anand Karandikar



Overview

- What is RDFa?
- RDFa Syntax and Attributes
 - Examples
- Chaining
- CURIE
- RDFa v/s Microformat

- What is GRDDL?
- GRDDL examples

- RDFa and GRDDL together



What is RDFa?

- **RDFa** (or Resource Description Framework - in - attributes) is a set of extensions to XHTML to express structured data.
- Makes use of XHTML attributes as well as it's own attributes to mark up human-readable data with machine-readable indicators for browsers and other programs to interpret.
- Underlying abstract representation is RDF



RDFa Attributes

- @about a URIorSafeCURIE, used for stating what the data is about (a 'subject' in RDF terminology);
- @property: a whitespace separated list of CURIEs, used for expressing relationships between a subject and some literal text (also a 'predicate');
- @resource: a URIorSafeCURIE for expressing the partner resource of a relationship that is not intended to be 'clickable' (also an 'object');
- @datatype: a CURIE representing a datatype, to express the datatype of a literal;
- @typeof : a whitespace separated list of CURIEs that indicate the RDF type(s) to associate with a subject.

Example: Using RDFa attributes

```
<html xmlns="http://www.w3.org/1999/xhtml"
xmlns:foaf="http://xmlns.com/foaf/0.1/"
xmlns:dc="http://purl.org/dc/elements/1.1/" >
<head>
  <title>My home-page</title>
  <meta property="dc:creator" content="Mark Birbeck" />
  <link rel="foaf:topic"
ref="http://www.formsPlayer.com/#us" />
</head>
<body>...</body>
</html>
```

Adopted from: <http://www.w3.org/TR/rdfa-syntax/>

Another example

```
<html
  xmlns="http://www.w3.org/1999/xhtml"
  xmlns:biblio="http://example.org/"
  xmlns:dc="http://purl.org/dc/elements/1.1/" >
  <head>
    <title>Books by Marco Pierre White</title>
  </head>
  <body>
    I think White's book '
    <span about="urn:ISBN:0091808189" typeof="biblio:book" property="dc:title">
    Canteen Cuisine
    </span>'
    has quite advanced stuff,. You might also like
    <span about="urn:ISBN:1596913614" typeof="biblio:book"
    property="dc:description">
    White's autobiography
    </span>.
  </body>
</html>
```

Chaining

- RDFa has the notion of [chaining] which aims to combine statements together in as intuitive a way as possible, so as avoid unnecessary repetition of mark-up

```
<div about="http://dbpedia.org/resource/Albert_Einstein">
```

```
  <span property="foaf:name">Albert Einstein</span>
```

```
  <span property="dbp:dateOfBirth" datatype="xsd:date">1879-03-14</span>
```

```
  <div rel="dbp:birthPlace"
```

```
    resource="http://dbpedia.org/resource/Germany">
```

```
      <span property="dbp:conventionalLongName">Federal Republic of Germany</span> </div>
```

```
</div>
```



Chaining continued..

- In the previous example we can see that an object resource ('Germany'), has become the subject for nested statements. This mark-up also illustrates the basic chaining pattern of 'A has a B has a C' (i.e., Einstein has a birth place of Germany, which has a long name of "Federal Republic of Germany").

CURIE and URI Processing

- Since RDFa is ultimately a means for transporting RDF, then a key concept is the *resource* and its manifestation as a URI.
- RDF deals with complete URIs (not relative paths), then when converting RDFa to triples, any relative URIs will need to be resolved relative to the base URI.
- CURIE = Compact URI
- A CURIE is a convenient way to represent a long URI, by replacing a leading section of the URI with a substitution token.

Curie Example

The full URI for Albert Einstein on DPPedia is:
http://dbpedia.org/resource/Albert_Einstein

Using CURIE,

```
<div xmlns:db="http://dbpedia.org/">  
  <div about="[db:resource/Albert_Einstein]">  
    ... </div>  
</div>
```

Converting a CURIE to a URI

Involves 3 steps

- Split the CURIE at the colon to obtain the prefix and the resource.
- Using the prefix and the current in-scope mappings, obtain the URI that the prefix maps to.
- Concatenate the mapped URI with the resource value, to obtain an absolute URI.

RDFa in the news

- MySpace
- Dreamweaver with RDFa extension
- Whitehouse.gov – third-party content on this site is licensed Creative Commons Attribution 3.0 License

```
<xhv:license
  rdf:resource="http://creativecommons.org/licenses/by/3.0/us/" />
<cc:attributionName
  xml:lang="en">Whitehouse.gov</cc:attributionName>
<cc:attributionURL rdf:resource="http://www.whitehouse.gov" />
```

- XHTML → RDF/XML
<http://www.w3.org/2007/08/pyRdfa/>



RDFa v/s Microformat

- XML namespaces --- Flat namespaces
- Reuses data models in RDF --- new formats require new data model
- Introduces new metadata attributes --- uses HTML/XHTML attributes
- W₃C initiative --- unofficial and ad hoc



GRDDL

- Gleaning Resource Descriptions from Dialects of Languages
- A technique for obtaining RDF data from XML documents and in particular XHTML pages.
- Authors can explicitly associate XSLT transformation procedures using *link* element in *head* of the document
- Alternatively, the information needed to obtain the transformation may be held in an associated metadata profile document or namespace document.



GRDDL transformations

- Anyone can write one
- A *standard transform library* has been provided that can extract RDF that's embedded directly in XML or HTML using `<rdf:RDF>` tags as well as extract any profile transformations.
- GRDDL transformations can be made for any dialect, even for microformats.

GRDDL example

- **Step 1**

Include profile URI for GRDDL in the document to declare that the markup can be interpreted using GRDDL.

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
  "http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">
  <head profile="http://www.w3.org/2003/g/data-view">
    <title>Robin's Schedule</title>
  </head>
  <body> ...
```




- **Step 2**

Add a *link* element containing the reference to the specific GRDDL transformation

(here HTML in hCalendar is transformed to RDF)

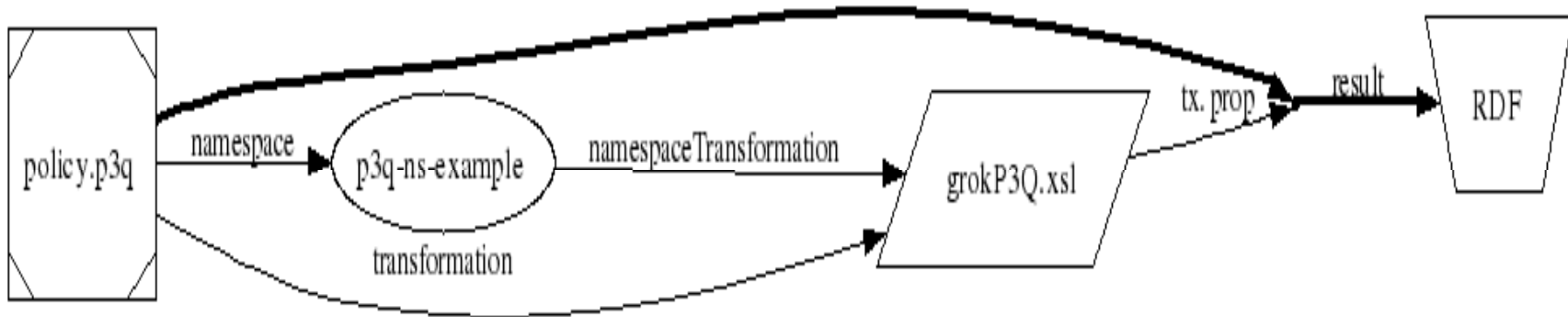
```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
  "http://www.w3.org/TR/xhtml1/DTD/xhtml1.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">
  <head profile="http://www.w3.org/2003/g/data-view">
    <title>Robin's Schedule</title>
    <link rel="transformation"
      href="http://www.w3.org/2002/12/cal/glean-hcal"/>
  </head>
  <body> .....
```



Using GRDDL with XML namespace

To associate a GRDDL transformation with a whole dialect, include a `grddl:namespaceTransformation` property in a GRDDL result of the namespace document.

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dataview="http://www.w3.org/2003/g/data-view#">
  <rdf:Description rdf:about="http://www.w3.org/2004/01/rdxh/p3q-ns-
    example">
    <dataview:namespaceTransformation
      rdf:resource="http://www.w3.org/2004/01/rdxh/grokP3Q.xsl"
    />
  </rdf:Description>
</rdf:RDF>
```



Every document whose root namespace name is ...p3q-ns-example has grokP3Q.xsl as a **GRDDL transformation** implicitly

Adopted from GRDDL Primer



Putting it together

- HTML for Jane's schedule
- RDFa Distiller gives the RDFa mark-up that the document contains
- GRDDL validator generates RDF
- Place RDF data in some RDF store

References

- http://rdfa.info/wiki/RDFa_Wiki
- <http://www.w3.org/TR/rdfa-syntax/>
- <http://en.wikipedia.org/wiki/RDFa>
- RDFa Distiller, <http://www.w3.org/2007/08/pyRdfa/>
- Examples adopted from GRDDL-primer
- GRDDL validator , <http://www.w3.org/2007/08/grddl/>



Questions