

## Metadata and Resource Description Framework (RDF)

Marian F Ursu  
Department of Computing, Goldsmiths College  
February 2003

## Resource

- resource = any object
  - any object that can be identified uniquely by a Uniform Resource Identifier (URI) plus an optional anchor
- our understanding associates some information/knowledge with each resource
  - it would be dangerous to say that this information/knowledge is inherent to the resource, because the subjective nature of understanding/interpreting would be lost
  - however, we could say that in many situations, there is an agreed understanding
  - without the above assumption, the metadata initiative would simply not be possible

## Metadata

- metadata = any textual description that expresses some of the knowledge/information inherent to a resource
  - resource = a picture
    - metadata - any statement describing that picture (its content, value, period, author, location, ...)
  - resource = a text
    - metadata - any statement describing that text ...

## Motivation for Metadata on the Web

- resources on the web = data objects
- data objects exist on the web in various formats
  - can be read by software
  - cannot be understood by software, unless the format is text
    - even in this case, the understanding is very limited
- metadata could be used to help software agents "understand" the data objects

## Types of Metadata Descriptions

- metadata can describe:
  - the content of a data object
    - summarise
    - enhance (e.g. provide context)
    - critique
    - constrain (if input is permitted)
  - the history of the data object
  - the access restrictions on the data object
  - links/relationships with other data objects

## Usage of Metadata

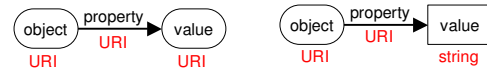
- content rating
- resource discovery
- resource integration
- communication between intelligent agents
- personalisation
  - metadata can describe (privacy) preferences of a user

## RDF

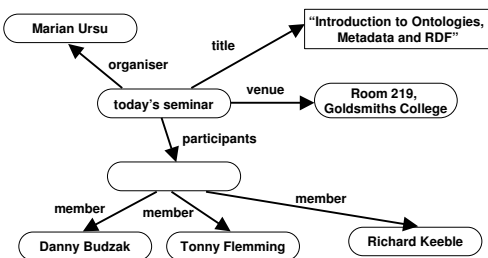
- domain independent (and syntax independent) model for resource descriptions
- aimed at supporting interoperability between software agents

## RDF

- essentially it is a triplet/statement
  - subject—predicate—object
  - object—property—value
 where each element of a triplet is a resource
  - exception: third element (value/object) can be a string
- graphical representation as a directed graph



## Example - Directed Graph



apart from the value of title, we should have used URIs

## Reasoning with RDF

- a software agent “understands” some data if it can “reason” with it
- RDF descriptions (metadata) are intended to be process-able by software agents
- RDF does not specify any mechanism for reasoning
- individual software agents would have to build such reasoning mechanisms

## Syntax: XML based RDF

- the syntax of RDF could be built on that of XML
  - choice motivated by the status of XML as a standard language for data interchange
  - other syntaxes for RDF may also be possible
- two types
  - serialisation
    - supports the full capabilities of the model in a regular fashion
    - cannot be included in HTML
  - abbreviated
    - more compact, achieved with more additional constructs
    - can be included in HTML
    - has limitations
  - the two can be mixed

## Simple - Example

```
<rdf:RDF>
  <rdf:Description about="http://www.mock/example/application.pdf">
    <schema:validity>2004-01-01</schema:validity>
    <schema:type>financial</schema:type>
  </rdf:Description>
</rdf:RDF>
```

- ‘Description’ element groups together a set of statements about a resource
- the resource is specified in the ‘about’ attribute
- property names must be associated with a **schema**
  - for now, assume that schema=metadata standard
  - this is how meaning is expressed

## Namespaces and Resources - Example

```
<?xml version 1.0>
<rdf:RDF
  xmlns:rdf="http://w3.org/TR/1999/PR-rdf-syntax-19990105#"
  xmlns:eGMS="http://mock.gov.uk/eGMS-syntax-20010301">
  <rdf:Description about="http://www.mock/example/application.pdf">
    <eGMS:date:valid>2004-01-01</eGMS:date:valid>
    <eGMS:subject>financial application form</eGMS:subject>
    <eGMS:creator rdf:resource="http://mock.gov.uk/marian-ursu"/>
  </rdf:Description>
</rdf:RDF>
```

- namespaces can be declared in RDF
  - default name spaces could be used
- values of properties may be resources
  - use the 'resource' attribute with that property

## Qualified/Complex Property Values - Example

```
<?xml version 1.0>
<rdf:RDF
  xmlns:rdf="... xmlns:eGMS="... xmlns:vCard="...
  <rdf:Description about="http://www.mock/example/application.pdf">
    <eGMS:date:valid>2004-01-01</eGMS:date:valid>
    <eGMS:subject>financial application form</eGMS:subject>
    <eGMS:creator rdf:resource="#creator01"/>
  </rdf:Description>
  <rdf:Description ID="#creator01">
    <vCard:name>Marian Ursu</vCard:name>
    <vCard:email>m.ursu@gold.ac.uk</vCard:email>
    <vCard:affiliation>Goldsmiths College</vCard:affiliation>
  </rdf:Description>
</rdf:RDF>
```

- shorthand: <eGMS:creator parseType=Resource><vCard:name>...

## Containers

- bag
  - unordered list of resources or literals
- sequence
  - ordered list of resources or literals
- alternative
  - a list of resources or literals that represent possible alternatives for the single value of a property
    - choose one from all of them
- to create a collection, RDF uses an additional resource
  - 'type' property could be 'Bag', 'Seq' and 'Alt'

## Statements about Statements

- reification
  - application—creator—marian is transformed into
    - statement(subject:application; property:creator; value:marian)
- the initial statement thus becomes a resource and statements can be made about it, in turn

## Examples

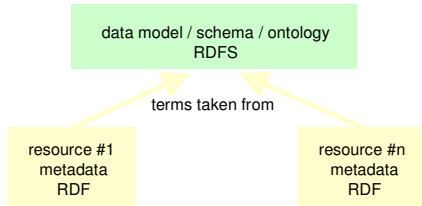
- <http://www.ukoln.ac.uk/metadata/rresources/rdf/examples/1>
- <http://www.ukoln.ac.uk/metadata/rresources/rdf/examples/2>
- <http://www.ukoln.ac.uk/metadata/rresources/rdf/examples/3>
- <http://www.ukoln.ac.uk/metadata/rresources/rdf/examples/4>

## Schema Definition - RDF Schema (RDFS)

- the specification of a metadata standard (or an ontology, or a vocabulary, or a data model) is called a schema
- RDFS is a language for such specifications
- in essence, an RDF Schema is similar to an ER model;
  - it is more verbose;
  - it is machine process-able
  - it uses different constraints
  - it has an OO flavour (classes + inheritance)
  - it supports references to other schemas
    - huge advantage in terms of interoperability
- other notations: XML Schemas

## Exercise

- compile a small ontology in NL and use some of its terms in resource description
  - focus on the automatic reasoning possible to achieve



## Further Details

<http://www.w3.org/TR/rdf-schema/>

## Conclusion

- RDF
  - a domain and syntax independent model for resource descriptions
  - recommended XML syntax
- RDFS
  - language for the definition of schemas (specification of ontologies or metadata vocabularies)
- probably, you will never have to use them directly

## References

- W3C, 1999. Resource Description Framework (RDF) Model and Syntax Specification. <http://www.w3.org/TR/PR-rdf-syntax/>
- W3C, 2002. RDF Vocabulary Description Language 1.0: RDF Schema. <http://www.w3.org/TR/rdf-schema/>
- Iannella, Renato, 1998. An Idiot's Guide to the Resource Description Framework. The New Review of Information Networking, Vol 4.
- Berners-Lee, Tim, xxx. Primer: Getting into RDF & Semantic Web using N3.
- Miller, Eric, 1998. An Introduction to the Resource Description Framework. D-Lib Magazine, May 1998.