An Ontology of Resources for Linked Data

Harry Halpin and Valentina Presutti

Advanced Semantic Technology, Spring 2009
Joshua Shinavier
Outline

• motivation
• modeling the Linked Data domain
• the IRW (Identity of Resources on the Web) ontology
• IJCAI review
Motivation

• clarify the (somewhat controversial) concepts of the Linked Data domain

• provide a framework for HTTP and Linked Data evaluation

• serve as a domain model for Linked Data tools

• allow Linked Data to describe itself
What is a resource?

- a resource is:
  - “a network data object or service, identified by a URI” (first HTTP RFC)
  - anything which might be identified by a URI (W3C TAG)
  - anything that has identity (Tim Berners-Lee, RFC 2396)
Web-accessibility of resources

• reference vs. access (Hayes)

  • **reference**: URI is used to mention the thing

  • **access**: URI provides a casual pathway to the thing

    • a resource is **Web-accessible** if it can be accessed via the use of HTTP

• resources need not be Web-accessible

  • e.g. human beings, the Eiffel Tower
Resources of Linked Data

- ability to name things which aren’t Web-accessible makes the Semantic Web, Linked Data possible

- Linked Data adds Web-accessible associated descriptions
A hierarchy of resources

• information resource
  • something “whose essential characteristics can be conveyed in a message” (W3C TAG)
  • e.g. web pages, multimedia files

• Web resource: an information resource which is also Web-accessible

• non-information resource
  • something which is not possibly Web-accessible
  • e.g. real people, abstract concepts like “the integers”

• Semantic Web resource: a non-information resource described in the Semantic Web
Linked Data and Redirection

- Linked data allows access of associated descriptions from URIs from non-information resource by use of redirection

- TAG’s official resolution:
  - **2xx**: response for GET of an information resource
  - **303**: response for GET of either an information or non-information resource
  - **4xx**: nature of resource is unknown
Figure 1: 303 Redirection for Semantic Web URIs
The hash URI alternative

• the hash convention is an alternative to 303 redirection

• e.g. <http://www.tour-eiffel.fr/#it>

• fragment identifier is ignored by the client

• URI before the hash identifies an information resource and is retrieved directly:
  • <http://www.tour-eiffel.fr/>

• neither technique answers the question of whether a resource is an IR or non-IR

• 303 response may redirect to another non-IR

• hash convention is dependent on media types
Why use an ontology?

- to describe Linked Data and typical Linked Data transactions
- existing ontologies lack coverage of the whole problem:
  - **HTTP in RDF** -- does not model the concept of resource
  - **RDF Schema** -- lacks IR / non-IR distinction
  - **Tabulator / AJAR** -- ?
  - **IRE** (Identifiers, Resources, and Entities) lacks distinction between resources and their Web representations, interaction between client and server
- **IRW** is its successor
The IRW OWL-DL ontology

- **Thing**
  - **Nothing**
  - **InformationObject = InformationResource**
- **InformationRealization**
  - **WebRepresentation**
  - **OntologyProperty**
- **Resource**
  - **InformationObject = InformationResource**
  - **InformationResource = InformationObject**
    - **WebResource**
    - **MediaType**
  - **NonInformationResource**
    - **AbstractResource**
    - **ConceptualResource**
  - **PhysicalEntityResource**
    - **InformationRealization**
      - **WebRepresentation**
- **URI**
  - **SemanticWebURI**
  - **WebClient**
  - **WebServer**
303 redirection in the IRW ontology
Identification and reference

- irw:URI
- allows hierarchy of URI types, including IRIs and Semantic Web URIs
- irw:hasURI
- irw:identifies, irw:isIdentifiedBy
- irw:accesses
- irw:refersTo, irw:isReferencedBy
Access and redirection

- irw:redirectsTo
- tag:redirects303To
- tag:redirectsHashTo
Types of resources

- irw:Resource
- irw:InformationResource
- irw:WebResource
- irw:NonInformationResource
- ir:InformationRealization
- irw:WebRepresentation
- ir:realizes, ir:isrealizedBy
- irw:isAbout
- irw:isTopicOf
More on non-information resources

• **irw:PhysicalEntityResource** -- “touchable”

• **irw:ConceptualResource** -- created in a social process

• **irw:AbstractResource** -- cannot be located in space-time
Hypertext Web transactions

- a irw:WebClient irw:requests a URI
- irw:WebServer
- irw:isResolution, irw:resolvesTo
- irw:isLocationOf
Linked Data transactions

- irw:SemanaticWebURI
- ldow:AssociatedDescription
### Mapping IRW to other ontologies

<table>
<thead>
<tr>
<th>Class or Property</th>
<th>Alignments</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>irw:WebRepresentation</code></td>
<td><code>owl:equivalentClass http:Message</code></td>
</tr>
<tr>
<td></td>
<td><code>owl:equivalentClass ont:ResponseMessage</code></td>
</tr>
<tr>
<td></td>
<td><code>rdfs:subClassOf ire:InformationRealization</code></td>
</tr>
<tr>
<td></td>
<td><code>rdfs:subClassOf ir:InformationRealization</code></td>
</tr>
<tr>
<td><code>http:Content</code></td>
<td><code>rdfs:subClassOf ir:InformationRealization</code></td>
</tr>
<tr>
<td><code>http:MessageHeader</code></td>
<td><code>rdfs:subClassOf ir:InformationRealization</code></td>
</tr>
<tr>
<td><code>irw:InformationResource</code></td>
<td><code>owl:equivalentClass ir:InformationObject</code></td>
</tr>
<tr>
<td><code>irw:SemanticWebURI</code></td>
<td><code>ire:SemanticWebURI</code></td>
</tr>
<tr>
<td><code>irw:identifies</code></td>
<td><code>ire:isExactProxyFor</code></td>
</tr>
<tr>
<td><code>irw:isAbout</code></td>
<td><code>ire:about</code></td>
</tr>
</tbody>
</table>
Applications of IRW

- make Linked Data **self-describing**
- embedding of an IRW statement:
  - in RDF documents, RDFa, GRDDL, SPARQL results, etc.
  - using the HTTP link header
- semantic HTTP **validation**
  - namespace validation using W3C’s EARL (Evaluation and Report Language)
- allow publishers to better describe their data with **Semantic Web Sitemaps**
- Linked Data **metadata**
- systematize Linked Data validation
Conclusion

- IRW:
  - is a first step towards modeling the Linked Data domain
  - clarifies the interactions between the hypertext Web and Linked Data
  - needs to be standardized
  - helps to grow the “Dark Side of the Semantic Web” (the Web side)
Review

• overall: 6.7 -- accept
• relevance: 9
• significance: 7
• technical soundness: 6
• novelty: 5
• usability: 5
• clarity: 8
Relevance (to a SW audience): 9

- deals with core concepts of Linked Data, an important sub-field of Semantic Web
- IR vs. non-IR, 303 redirection vs. hash URIs, etc. have been discussed extensively on the Web and on Semantic Web mailing lists, less so in academic literature
Significance: 7

- first ontology which can be used both for modeling HTTP interactions and relationships between Semantic Web and other resources
- may be useful for Semantic Web clients, e.g. for caching metadata
Technical soundness: 6

- Linked Data concepts (and controversies) are accurately represented
- Some choices of terminology are odd
- Ontology also slightly odd (e.g. property range of owl:Thing intersected with irw:Resource, which is a subclass)
- Discrepancies between ontology and examples
- IRW ontology not properly published as Linked Data
Novelty: 5

- tackles a well-known problem
- there is much related work
- greater coverage of the Linked Data domain than other ontologies in this area
Usability: 5

- use cases / applications are convincing
- controversial nature of the concepts of the ontology is an issue
- numerous minor errors and omissions
- why model a web client?
- modeling of non-information resources is perhaps too fine-grained
- authors point out that the ontology is only a first step and needs to be refined
Clarity: 8

- paper is fairly clearly written, easy to read and understand
- problem statement and solution are clearly presented
- the ontology is available on the Web, and is thoroughly annotated
- writing style is confusing in places
• high clarity of the paper and unambiguous nature of the ontology make this work relatively easy to evaluate

• reviewer is not an expert in the domain, but has a fair amount of experience

• some concepts from the paper (e.g. EARL, DUL +IOL) are unfamiliar to the reviewer
Questions?