Open Archives Initiative
Object Re-use & Exchange

Abstract Data Model and Implementation: Beyond the Basics

Simeon Warner
Pete Johnston, Carl Lagoze, Michael Nelson, Robert Sanderson, Herbert Van de Sompel

(1) simeon@cs.cornell.edu

Open Repositories 2008
Southampton University, UK. 4 April 2008
1. Proxies

Adding context to resources
Alice and Bob observe cats in Eve’s laboratory

ReM-1 \(\xrightarrow{\text{ore:describes}}\) A-1 \(\xrightarrow{\text{ore:aggregates}}\) AR-1

AR-1 \(\xleftarrow{\text{hasNext}}\) AR-2

AR-2 \(\xleftarrow{\text{hasNext}}\) AR-3

Cat observations in Eve’s laboratory

Observation by Alice on 2008-04-01

Observation by Bob on 2008-04-02

Observation by Alice on 2008-04-03
Alice’s observations

- ReM-2 ore:describes A-2 ore:aggregates AR-3
- Cat observations by Alice
- Observation by Alice on 2008-04-01
- hasNext
- Observation by Alice on 2008-04-03
Agent combines ReM-1 and ReM-2 and is confused

Observation by Alice on 2008−04−01
Observation by Bob on 2008−04−02
Observation by Alice on 2008−04−03

hasNext

AR−1

hasNext

AR−2

hasNext

AR−3

Observation by Alice on 2008−04−01
Observation by Bob on 2008−04−02
Observation by Alice on 2008−04−03
What did we mean by hasNext?

ReM-1 — Bob’s observation on 2008-04-02 is the next observation after Alice’s observation on 2008-04-01 in the sequence of observations in Eve’s laboratory

ReM-2 — Alice’s observation on 2008-04-03 is the next observation after her observation on 2008-04-01 in the sequence of Alice’s observations
Modelling a resource in context

Two components:

- the **Resource**
- the context, which in ORE is the **Aggregation**

In web architecture a new concept needs a new resource (and hence name/identifier)… enter the **Proxy**:

![Diagram](image)
Alice’s next observation in context

Cat observations by Alice

ore:describes ore:aggregates

Observation by Alice on 2008-04-01

Observation by Alice on 2008-04-03

ore:proxyIn
ore:proxyFor

ore:proxyIn
ore:proxyFor

AR-1
P-1
A-2
ReM-2
AR-3
P-1
A-2
ReM-2

Observation by Alice on 2008-04-01

hasNext
2. Lineage

A resource may be in multiple Aggregations, how can we indicate provenance?

How can we say “I got it here”? 
Lineage depends on context

I (recipient context)
got it (resource in both contexts)
here (original context)

In ORE proxies provide context.

⇒use ore:lineage as relation between proxy nodes
Resource Maps with Lineage
3. **Nesting** Nesting or tree structures are common:

- Repository, Collection, Item, (Version, Datastream,...)
- Journal, Issue, Article (Version, Format)
- Artist, Album, Track
- Collaboration, Experiment, Result set, Data file, Data segment
No nested Aggregations in a Resource Map

The ORE Data Model includes only the notion of a single level of aggregation. How do I represent aggregations of aggregations?
Structure within an Aggregation

ReM-1 \(\text{ore:describes}\) A-1

\(\text{ore:aggregates}\)

A-1 \(\text{dcterms:isPartOf}\) AR-1
A-1 \(\text{dcterms:isPartOf}\) AR-2
A-1 \(\text{dcterms:isPartOf}\) AR-3

AR-1 \(\text{dcterms:isPartOf}\) AR-1a
AR-1 \(\text{dcterms:isPartOf}\) AR-1b
AR-2 \(\text{dcterms:isPartOf}\) AR-2a
AR-2 \(\text{dcterms:isPartOf}\) AR-2b
AR-3 \(\text{dcterms:isPartOf}\) AR-3a
AR-3 \(\text{dcterms:isPartOf}\) AR-3b
Nested Aggregations
4. RDF/XML and RDFa

- The ORE Data Model is defined in RDF
- Serialize in RDF/XML or RDFa by feeding the RDF for a Resource Map to any compliant library?

Atom vs RDF Syntaxes

pro Atom

• Widely used and understood
• Good tools and libraries
• ORE/Atom specific libraries being developed
• RDF/XML can be generated via GRDDL

pro RDF Syntaxes

• Complete expressive power in one style
• No mapping required
• Easy extensibility
• RDFa can be embedded in XHTML
RDF/XML structure and serialization
RDF/XML Syntax Specification (Revised)

<rdf:Description rdf:about="http://www.w3.org/TR/rdf-syntax-grammar">
  <ex:editor>
    <rdf:Description>
      <ex:homePage>
        <rdf:Description rdf:about="http://purl.org/net/dajobe/">
          </rdf:Description>
      </ex:homePage>
      <ex:fullName>Dave Beckett</ex:fullName>
    </rdf:Description>
  </ex:editor>
  <dc:title>RDF/XML Syntax Specification (Revised)</dc:title>
</rdf:Description>

...
RDF/XML single level XML — ORE style

...<rdf:Description rdf:about="http://www.w3.org/TR/rdf-syntax-grammar">
  <ex:editor rdf:resource="_blank" />
</rdf:Description>

<rdf:Description rdf:about="_blank">
  <ex:homePage rdf:resource="http://purl.org/net/dajobe/" />
  <ex:fullName>Dave Beckett</ex:fullName>
</rdf:Description>

<rdf:Description rdf:about="http://www.w3.org/TR/rdf-syntax-grammar">
  <dc:title>RDF/XML Syntax Specification (Revised)</dc:title>
</rdf:Description>

...
ORE Resource Map RDF/XML profile

Why create a profile?

- Make data amenable to processing with more tools

Rules:

- One `rdf:Description` for each distinct subject in Resource Map (`rdf:about subject`)
- Each predicate is element within `rdf:Description` following usual QName mapping
- Resource object: `rdf:resource` attribute, empty element
- Literal object: content of element (also typing rules)
- Can use `xml:base` and relative URIs
xml:base="http://dlib.org/dlib/february06/smith/02smith/">

<rdf:Description rdf:about="rem">
  <ore:describes rdf:resource="agg" />
  <dc:creator rdf:resource="http://example.org/AgencyX" />
  <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2008-02-12</dcterms:modified>
  <dc:rights rdf:resource="http://example.org/docs/doc123/" />
</rdf:Description>

<rdf:Description rdf:about="agg">
  <rdf:type rdf:resource="http://www.openarchives.org/ore/terms/Aggregation" />
  <ore:aggregates rdf:resource="http://example.org/docs/doc456/" />
  <ore:aggregates rdf:resource="http://example.org/docs/doc457/" />
</rdf:Description>

<rdf:Description rdf:about="http://example.org/docs/doc456/">
</rdf:Description>
Resource Maps in RDFa

- Embed Resource Map in XHTML page
- Splash page can do double duty as machine and human entry point
- Set of rules for ORE use to promote interoperability
  - base is Resource Map URI (via document URI, xml:base or base)
  - RDF data via RDFa attributes (about, resource, href, src, property, datatype, instanceof, rel and rev)
  - Some attributes support prefix:local-part CURIE syntax (extends QName)
5. HTTP implementation

- The way the web is built
- Aggregation and Resource Map have separate URIs
- Want to cite Aggregation
  \[\Rightarrow\text{need to be able to find Resource Map from Aggregation}\]
  \[\Rightarrow\text{RULE: must be a mechanism to do this}\]
**Cool URIs**

<table>
<thead>
<tr>
<th>Aggregation</th>
<th>A-1</th>
<th><a href="http://example.org/foo">http://example.org/foo</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Resource Map</td>
<td>ReM-1</td>
<td><a href="http://example.org/foo.xml">http://example.org/foo.xml</a></td>
</tr>
</tbody>
</table>

How do we get from Aggregation to Resource Map?

**Content Negotiation** — ReM-1 URI in Content-Location

**Redirection** — 303 ala Linked Data

<table>
<thead>
<tr>
<th>Aggregation</th>
<th>A-1</th>
<th><a href="http://example.org/foo">http://example.org/foo</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Resource Map</td>
<td>ReM-1</td>
<td><a href="http://example.org/foo.xml">http://example.org/foo.xml</a> (Atom)</td>
</tr>
<tr>
<td></td>
<td>ReM-2</td>
<td><a href="http://example.org/foo.rdf">http://example.org/foo.rdf</a> (RDF/XML)</td>
</tr>
</tbody>
</table>
No server support

If there is no server support available then may use URI fragment identifiers to “lead to” the Resource Map from either ReM-1 or A-1.

<table>
<thead>
<tr>
<th>Aggregation</th>
<th>A-1</th>
<th><a href="http://example.org/foo.xml">http://example.org/foo.xml</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Resource Map</td>
<td>ReM-1</td>
<td><a href="http://example.org/foo.xml#rem">http://example.org/foo.xml#rem</a></td>
</tr>
</tbody>
</table>

WARNING — still some questions about whether this is the best approach to recommend!
## RDFa

<table>
<thead>
<tr>
<th>Aggregation</th>
<th>A-1</th>
<th><a href="http://example.org/foo">http://example.org/foo</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Resource Map</td>
<td>ReM-1</td>
<td><a href="http://example.org/foo.html">http://example.org/foo.html</a> (+RDFa)</td>
</tr>
<tr>
<td></td>
<td>ReM-2</td>
<td><a href="http://example.org/foo.xml">http://example.org/foo.xml</a> (Atom)</td>
</tr>
<tr>
<td></td>
<td>ReM-3</td>
<td><a href="http://example.org/foo.rdf">http://example.org/foo.rdf</a> (RDF/XML)</td>
</tr>
</tbody>
</table>

Lead to HTML by default
6. Tools

- **Atom Feed validator** — general purpose validator for Atom feed documents at http://validator.w3.org/feed/check.cgi (check ex4.2)
  
  Can install locally and libraries available to automate use (e.g. WebService::Validator::Feed::W3C in Perl)

More tools

7. Where to start...

Open Archives Initiative
Object Reuse and Exchange

ORE Specification and User Guide - Table of Contents
21 March 2008

Note: This document is alpha and subject to change at any time. It is being made available to the public for review and comment. Any implementation of the specifications or recommendations within should be undertaken with recognition of this alpha status. Please comment via the OAI-ORE Google Group.

This version:
http://www.openarchives.org/ore/0.3/toc

Latest version:
http://www.openarchives.org/ore/toc

Previous version:
http://www.openarchives.org/ore/0.2/toc
... (and where I end)

- ORE User Guide - Primer for summary.
- Validators already described.
- OAI-ORE Google Group (oai-ore) for comments and discussion. We’d love additional feedback in this alpha phase and on through beta.
That’s all folks...
That's all folks...