Open Archives Initiative
Object Reuse & Exchange

Abstract Data Model and Serialization

Carl Lagoze (1)
Pete Johnston, Michael Nelson, Robert Sanderson, Herbert Van de Sompel, Simeon Warner

(1)Computing and Information Science
Cornell University

lagoze@cs.cornell.edu
http://www.cs.cornell.edu/lagoze
This Presentation

• OAI-ORE for the most of us
  – Abstract Data Model Basics
    • Aggregations
    • Resource Maps
    • Metadata
    • Aggregated Resource
    • Associating with similar resources
    • Relationships to other Resources
  – Using Atom to describe Aggregations
Document Chain

- Primer
- Resource Map Implementation in Atom
- Abstract Data Model
- Resource Map Profile of Atom
Abstract Data Model
Why and What?

• Separation of concerns
  – Design
  – Implementation

• Provide basis for future implementations
  – Technology of the web (e.g. HTTP) will change over time
  – Other implementations are possible
Requirements of the Model

• Aggregations
  – Both simple hierarchical and inter/intra related
  – Identification via URI
  – Metadata

• Resource Maps
  – Description of aggregations via a set of assertions
  – Identification via URI – independent of aggregation
  – Metadata

• Conformance to web architecture and RDF Semantics
So, the tools we have to solve the problem are:

- Resource
- URI
- Representation
Simple Aggregation

Describe an HTML page, PDF, and Latex as an aggregation

Hierarchical or *Tree* Structure
Hierarchical Models and XML
More Complexity

Describe an HTML page, PDF, and LaTeX as an aggregation.

Assert that the PDF and LaTeX are journal articles

*Graph* Structure, Typed Relationships
### Resource Description Framework (RDF)

**Triples**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Predicate</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>hasChapter</td>
<td>R2</td>
</tr>
<tr>
<td>R1</td>
<td>hasChapter</td>
<td>R3</td>
</tr>
<tr>
<td>R3</td>
<td>follows</td>
<td>R2</td>
</tr>
<tr>
<td>R1</td>
<td>createdBy</td>
<td>“Carl Lagoze”</td>
</tr>
</tbody>
</table>
Multiple serializations
ORE Specification - Abstract Data Model
28 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/datamodel
Latest version:
http://www.openarchives.org/ore/datamodel
Previous version:
http://www.openarchives.org/ore/0.2/datamodel

Editors (OAI Executive)
Carl Lagoze, Cornell University Information Science
Herbert Van de Sompel, Los Alamos National Laboratory

Editors (ORE Technical Committee)
Pete Johnston, Eduserv Foundation
Michael Nelson, Old Dominion University
Robert Sanderson, University of Liverpool
Simeon Warner, Cornell University Information Science

Abstract

Open Archives Initiative Object Reuse and Exchange (OAI-ORE) defines standards for the description and exchange of aggregations of Web resources. This document describes the abstract data model that is the foundation for these standards. This model is conformant with the Architecture of the World Wide Web [Web Architecture] and leverages Named Graphs
The starting point: bringing some resources together

These resources have URIs

This has no URI

The resource have representations

HTTP GET

HTTP GET

HTTP GET
Aggregation: Resource that is a set of resources

This resource is an Aggregated Resource

A-1

ore:aggregates

This resource is an Aggregation

AR-1

AR-2

AR-3
An Aggregated Resource is just a Resource
Describe this Aggregation: Resource Map

This resource is a Resource Map

HTTP GET

The resource has a representation
Relationship between Aggregation and Resource Map

- An Aggregation is a Resource with a URI
- A Resource Map is a Resource with a URI
- Each Resource Map asserts (identifies) and describes one Aggregation
- Each Aggregation MAY be asserted and described by multiple Resource Maps
- Each Resource MUST have one serialization (representation)
Relationship between Aggregation and Resource Map
Relationship between Aggregation and Resource Map

Diagram showing the relationship between URI-A, URI-R, and URI-1 with the concepts of describes and describedBy.
Trust and Authority

• Same as any metadata on the Web
• Addressed somewhat by discovery mechanisms (Aggregation can lead client to “authoritative” Resource Map)
Metadata about the Resource Map & Aggregation

The ReM makes the assertions

ReM-1

ore:describes

A-1

Metadata about the ReM

dct:modified

dc:creator

Required

Metadata about the Aggregation

dc:creator
Asserting Similarity to other Resources

The ReM makes the assertions

This is NOT the Aggregation. But the Resource Map MAY assert similarity to it

ore:describes

ore:simlarTo

ORE Open Meeting, OR’08, Southampton, UK
April 4, 2008
Asserting other Relationships

The ReM makes the assertions

Assertions about the Aggregation.

Aggregation is a journal

Aggregation has another version “A”

“AR-3” is by Stephen Hawking

Aggregated Resources are articles

ORE Open Meeting, ORe08, Southampton, UK
April 4, 2008
Limits of Assertions thus Far

• The meaning of an RDF triple is independent of the context in which it is stated

• Think of the difference:
  – Carl is a man
  – Carl is staying at Jury’s

• All the triples described thus far are context independent
  – Therefore they can have the URI of an aggregated resource as subject or object
  – But remember that is just the URI of the Resource and is not exclusive of it being an Aggregated Resource

• Stay tuned for more
Aggregations can be Complex
But most of the time they look like this….
Serializing ORE Instances
General Serialization Goals

• Express as much of the model as possible
• Ability to round-trip
  – serialization -> model triples -> serialization
• Use well-known standardized technologies
  – Leverage tools and knowledge
Syndication

- Simple
- Compound Document

RSS
- 0.x
- 1.0
- 2.0

Atom

OAI Object Reuse & Exchange: Basics
ORE Open Meeting, OR’08, Southampton, UK
April 4, 2008
Atom

- Attempt to rationalize RSS 1.x, 2.x divergence
- Encoding is up-to-date with current XML standards
  - namespaces
  - Schema
- Robust content model
  - Distinguishes between metadata and content (plain text, HTML, base-64 binary)
- Well-defined extensibility model
- IETF FRC 4287
  - http://www.ietf.org/rfc/rfc4287
Structure of Feed

- Feed
  - ID
  - Author
  - Link
  - Title
  - Updated
  - *

- Entry
  - ID
  - Updated
  - Link
  - Summary
  - Content
  - *
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns="http://www.w3.org/2005/Atom">
  <title>Dan’s Blog</title>
  <link href="http://netzoooid.com/blog/"/>
  <updated>2007-11-07T18:30:02Z</updated>
  <author>
    <name>Dan Diephouse</name>
  </author>
  <id>urn:uuid:60a76c80-d399-11d9-b91c-0003939e0af6</id>
  <entry>
    <title>Building services with AtomPub</title>
    <link href="http://netzoooid.com/blog/atompub_services"/>
    <id>urn:uuid:1225c695-cfb8-4ebb-aaaa-80da344efa6a</id>
    <updated>2007-11-07T18:30:02Z</updated>
    <content>(you must have content or a summary)</content>
  </entry>
</feed>
ORE Specification - Resource Map Profile of Atom
26 February 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/atom

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

Previous version:
http://www.openarchives.org/ore/0.2/atom

Editors (OAI Executive)
Carl Lagoze, Cornell University Information Science
Herbert Van de Sompel, Los Alamos National Laboratory

Editors (ORE Technical Committee)
Pete Johnston, Eduserv Foundation
Michael Nelson, Old Dominion University
Robert Sanderson, University of Liverpool
Simeon Warner, Cornell University Information Science

Abstract

Open Archives Initiative Object Reuse and Exchange (OAI-ORE) defines standards for the description
and exchange of aggregations of Web resources, named Aggregations. OAI-ORE introduces the notion
of Resource Maps that describe an Aggregation. A Resource Map is resource that is a specialization of
a named graph. A Resource Map identifies an Aggregation, it asserts the finite set of constituent
resources (the Aggregated Resources) of the Aggregation, and it can express types and relationship
Atom Serialization Goals

• Result should be valid Atom
• Use as many Atom constructs as possible
  – Don’t rely on extensions
  – Don’t distort semantics
• Rely on RDF/XML when necessary
## Mapping ORE Model to Atom

<table>
<thead>
<tr>
<th>anoORE</th>
<th>Atom</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Aggregation</strong></td>
<td><strong>Feed</strong></td>
</tr>
<tr>
<td>URI-A</td>
<td>Feed &lt;id&gt;</td>
</tr>
<tr>
<td>URI-R</td>
<td>&lt;link href=&quot;URI&quot; rel=&quot;self&quot;</td>
</tr>
<tr>
<td>ore:similarTo</td>
<td>&lt;link href=&quot;URI&quot; rel=&quot;related&quot;</td>
</tr>
<tr>
<td><strong>Aggregation Properties/Metadata</strong></td>
<td>Feed metadata</td>
</tr>
<tr>
<td><strong>Aggregated Resource</strong></td>
<td><strong>Entry</strong></td>
</tr>
<tr>
<td>URI-AR</td>
<td>&lt;link href=&quot;URI&quot; rel=&quot;alternate&quot;</td>
</tr>
<tr>
<td>Aggregated Resource Properties/Metadata</td>
<td>Entry metadata</td>
</tr>
</tbody>
</table>
Atom/ORE Skeleton

```xml
<?xml version="1.0" encoding="UTF-8"?>
<feed xmlns="http://www.w3.org/2005/Atom">
  <updated>2007-10-10T18:30:02Z</updated>
  <category scheme="http://www.openarchives.org/ore/terms/"
            term="http://www.openarchives.org/ore/terms/Aggregation" label="Aggregation" />
</feed>
```

This is an aggregation
Adding Entries

<?xml version="1.0" encoding="UTF-8"?>
<feed xmlns="http://www.w3.org/2005/Atom">
  <updated>2007-10-10T18:30:02Z</updated>
  <category scheme="http://www.openarchives.org/ore/terms/" />
  <entry>
  </entry>
  <entry>
  </entry>
  <entry>
  </entry>
</feed>
Data Model <-> Atom
Looking to the future with Atom

- Atom Publishing Protocol
- SWORD
- Microsoft/Google APIs