

Open Archives Initiative Object Re-use & Exchange

Abstract Data Model and Implementation: Beyond the Basics

Simeon Warner⁽¹⁾

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

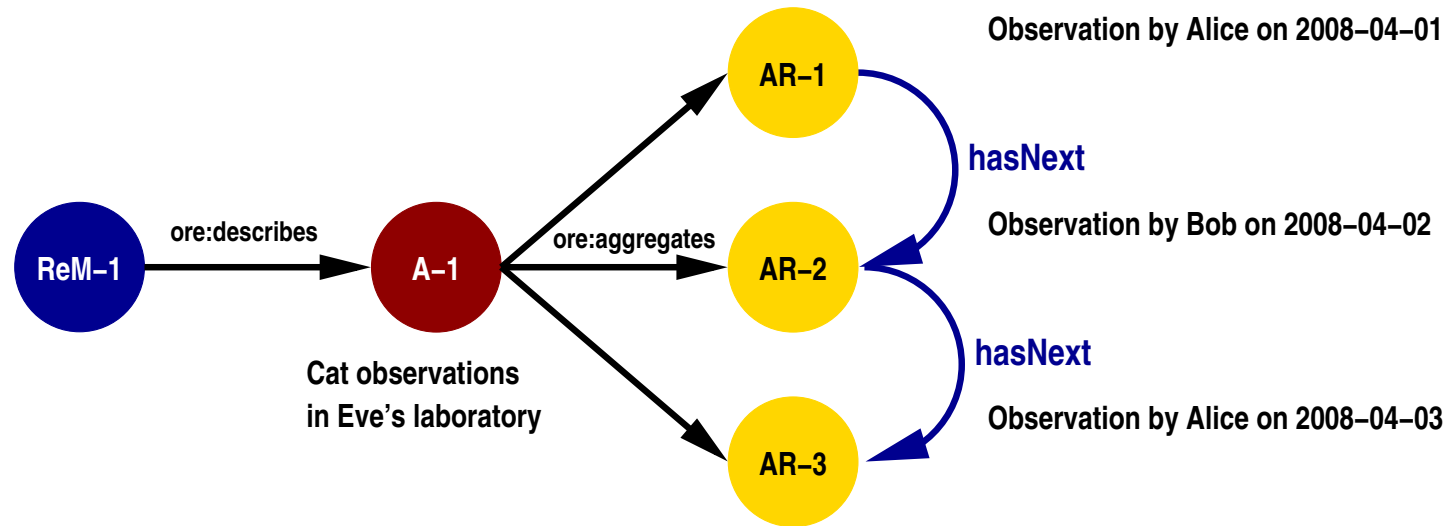
⁽¹⁾ simeon@cs.cornell.edu

Open Repositories 2008
Southampton University, UK. 4 April 2008

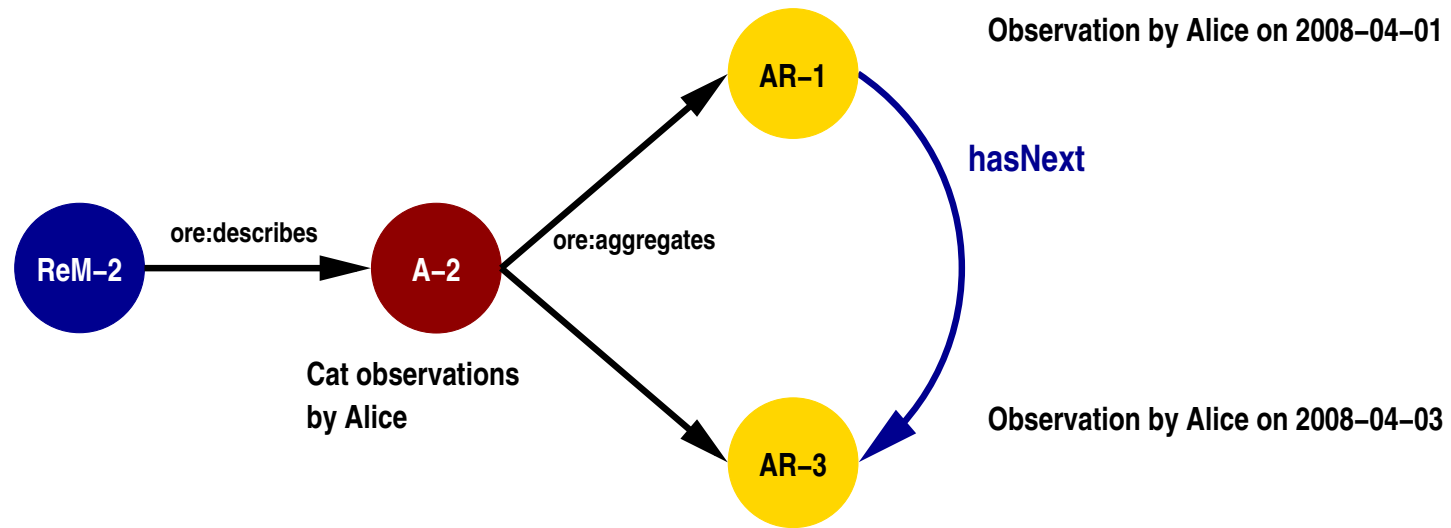
1. Proxies

Adding *context* to resources

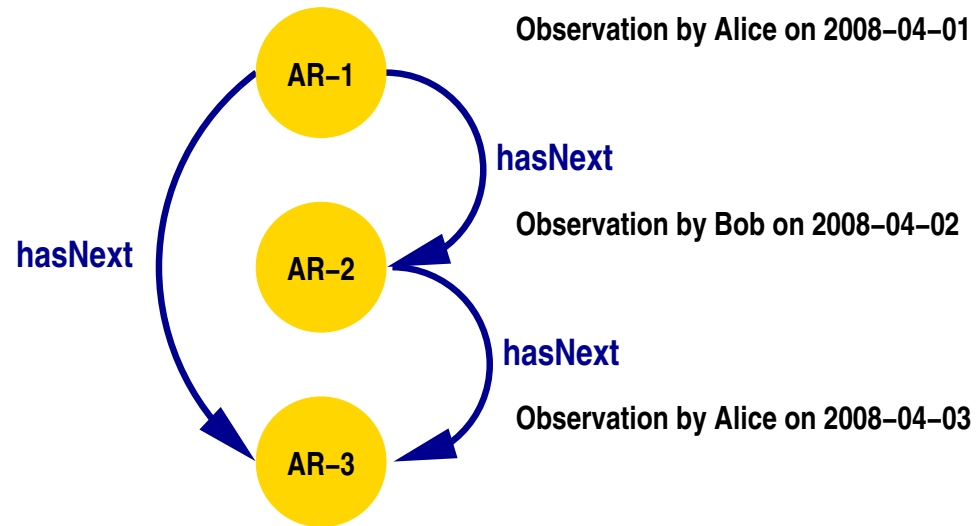
Alice and Bob observe cats in Eve's laboratory



Alice's observations



Agent combines ReM-1 and ReM-2 and is confused



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*

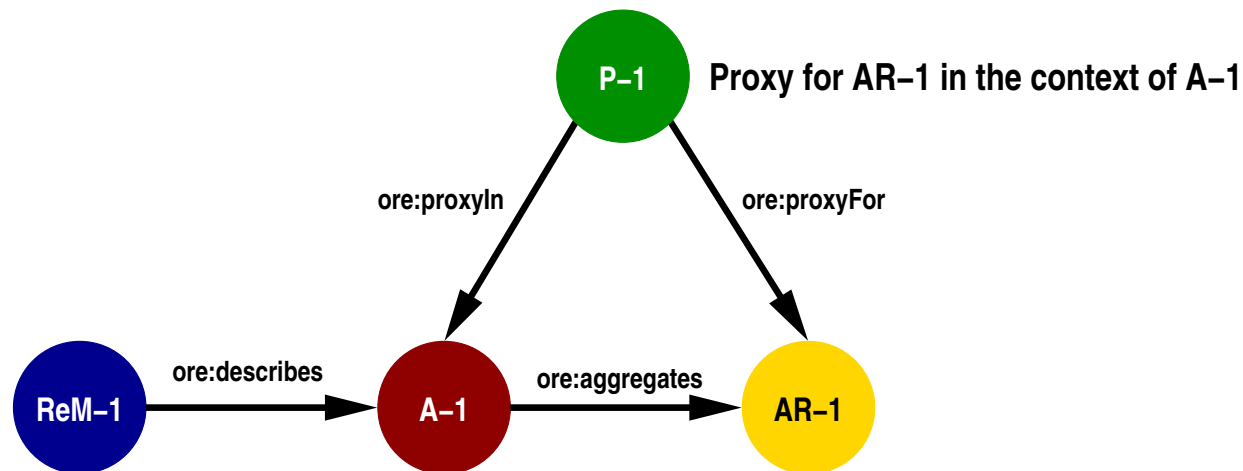
ReM-2 — Alice's observation on 2008-04-03 is the next observation after her observation on 2008-04-01 *in the sequence of Alice's observations*

Modelling a resource in context

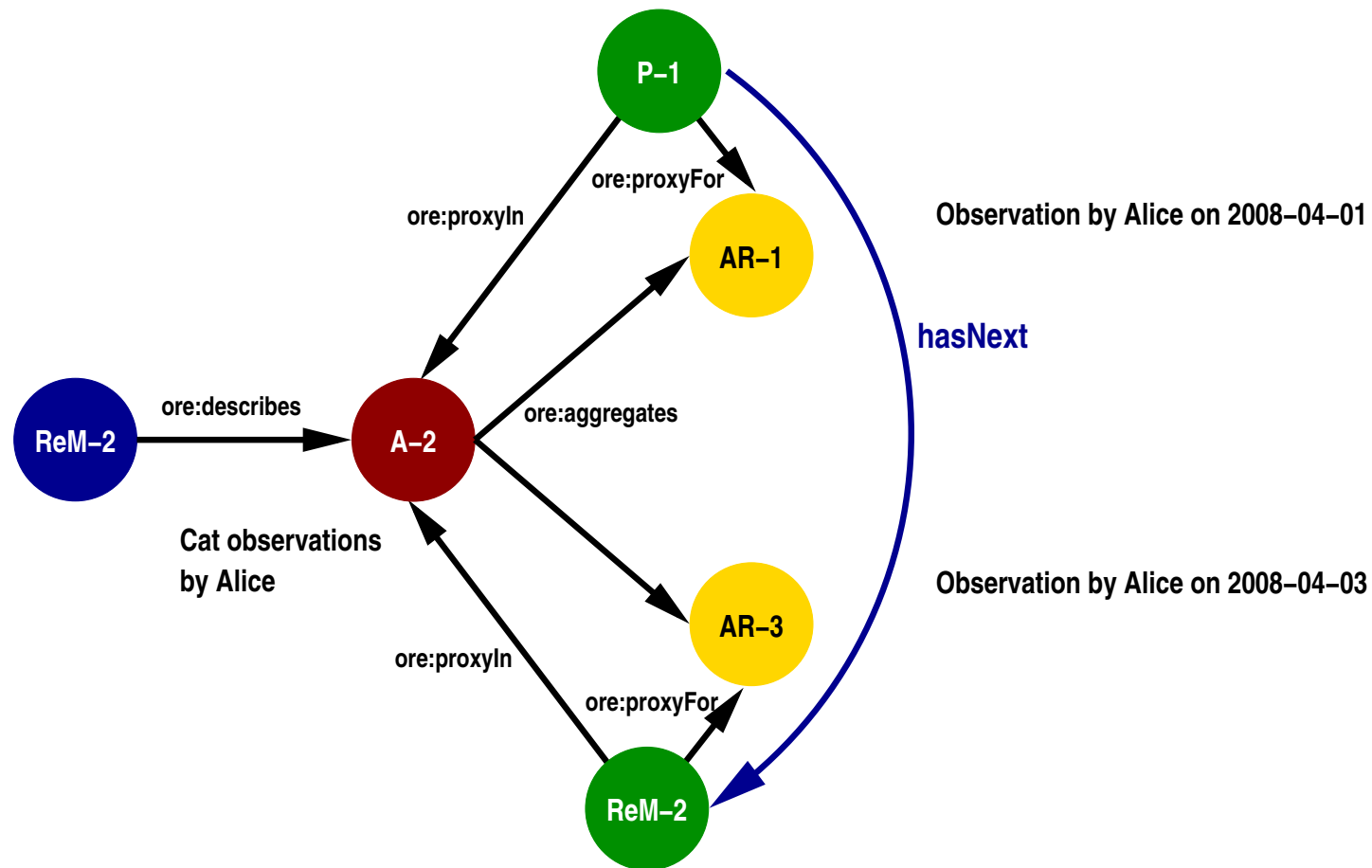
Two components:

- 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**:



Alice's next observation in context



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

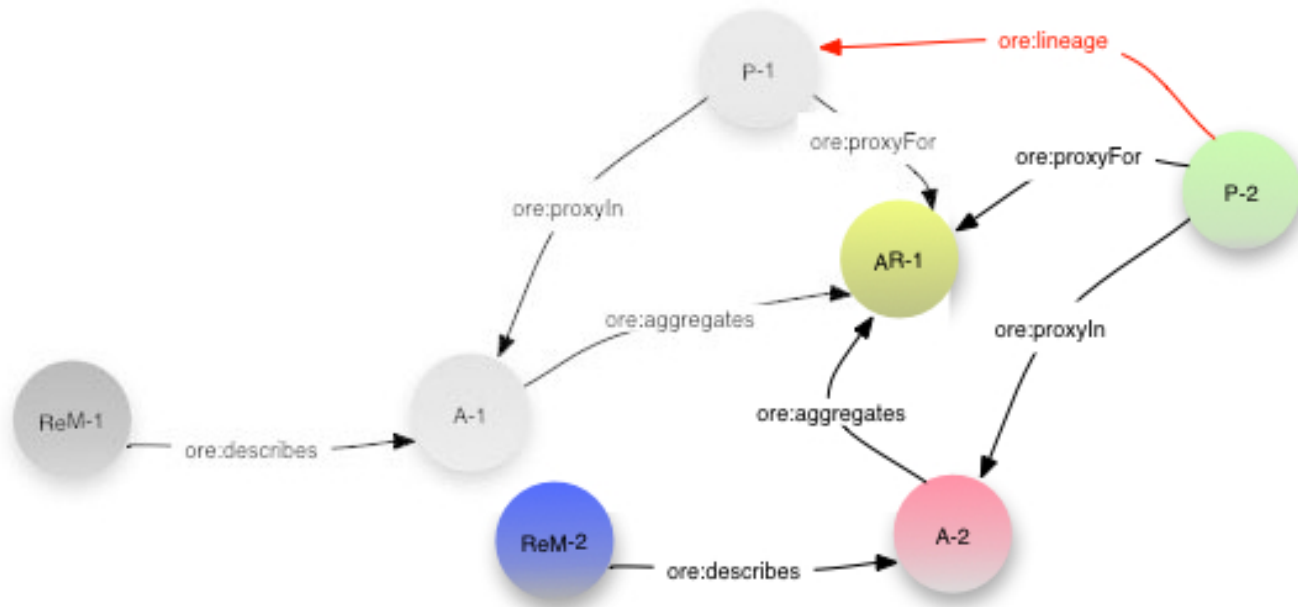
it (resource in both contexts)

here (original context)

In ORE *proxies* provide context.

⇒ use ore:lineage as relation between proxy nodes

Resource Maps with Lineage



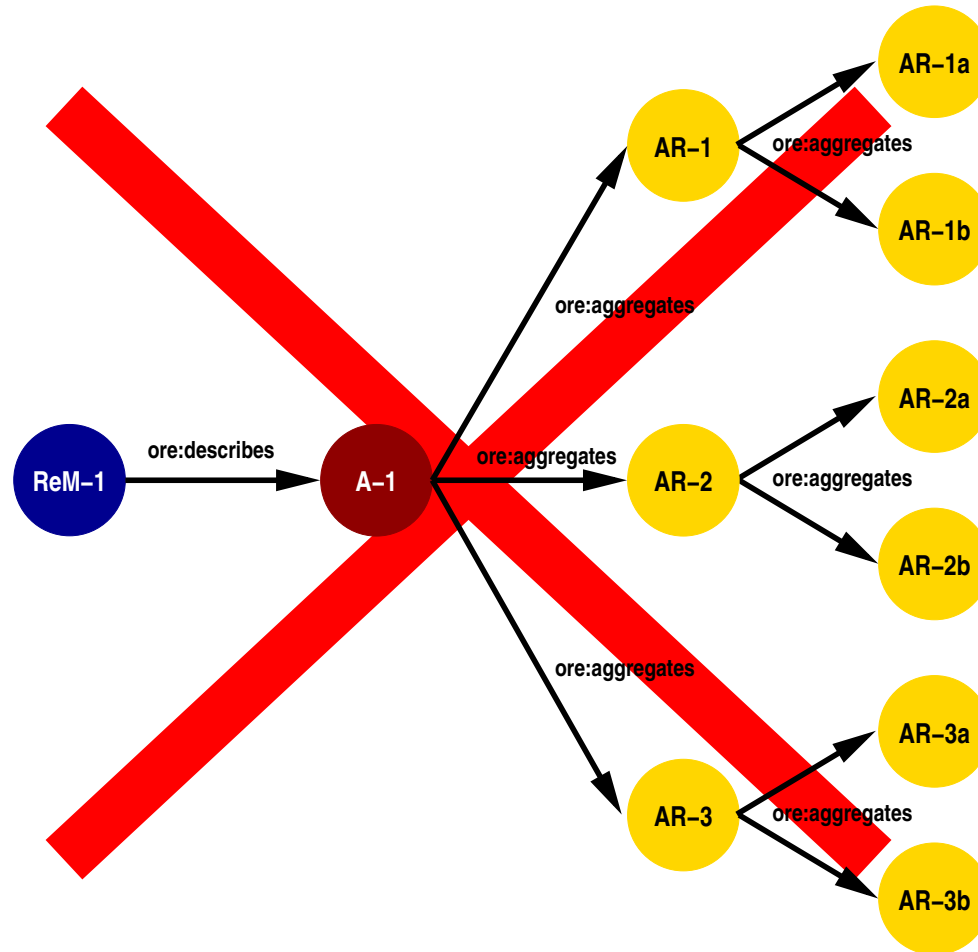
<P-2>	<ore:lineage>	<P-1>	ore:lineage	http://www.openarchives.org/ore/terms/lineage
-------	---------------	-------	-------------	---

3. Nesting Nesting or tree structures are common:

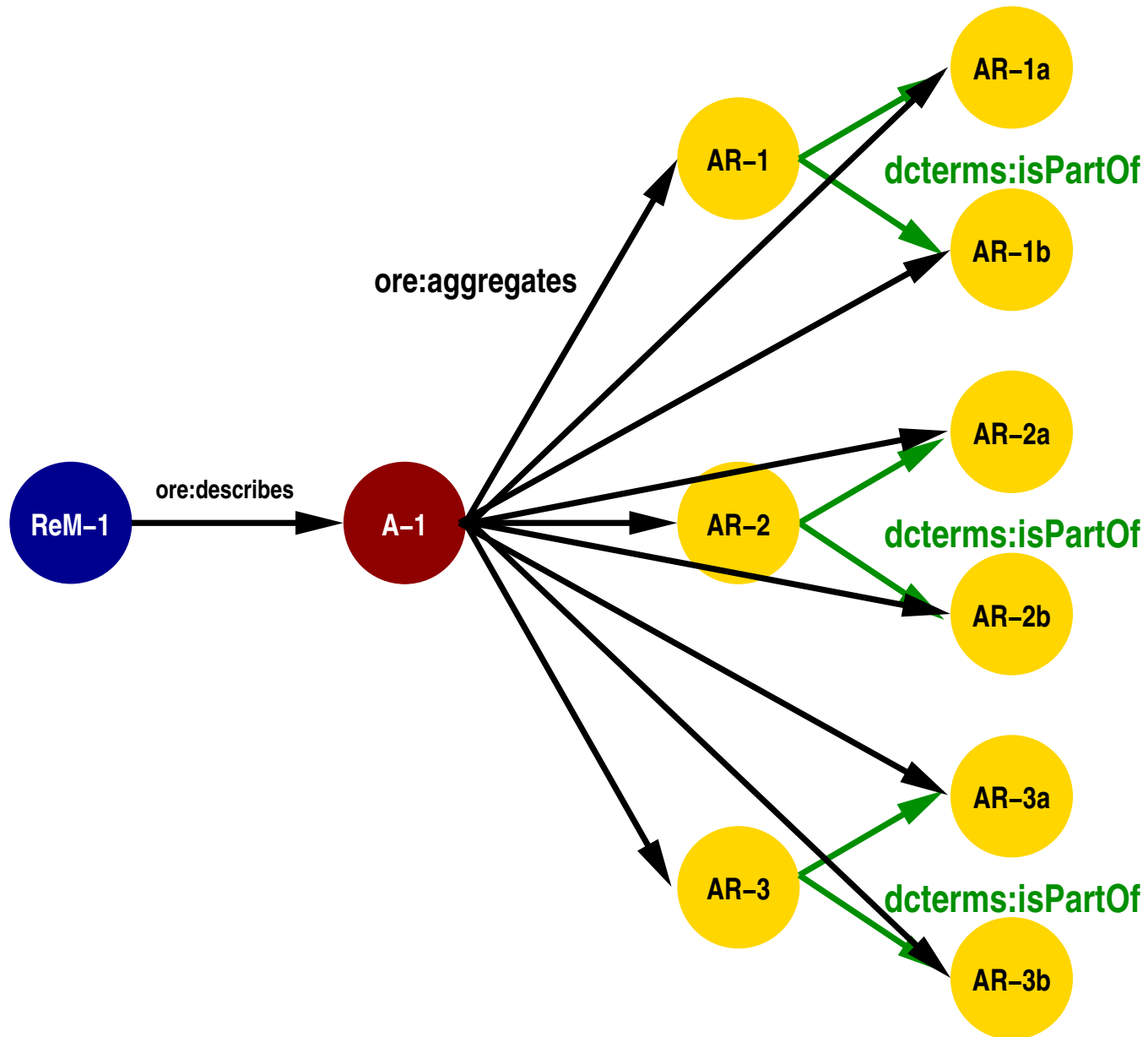
- Repository, Collection, Item, (Version, Datastream,...)
- Journal, Issue, Article (Version, Format)
- Artist, Album, Track
- Collaboration, Experiment, Result set, Data file, Data segment

No nested Aggregations in a Resource Map

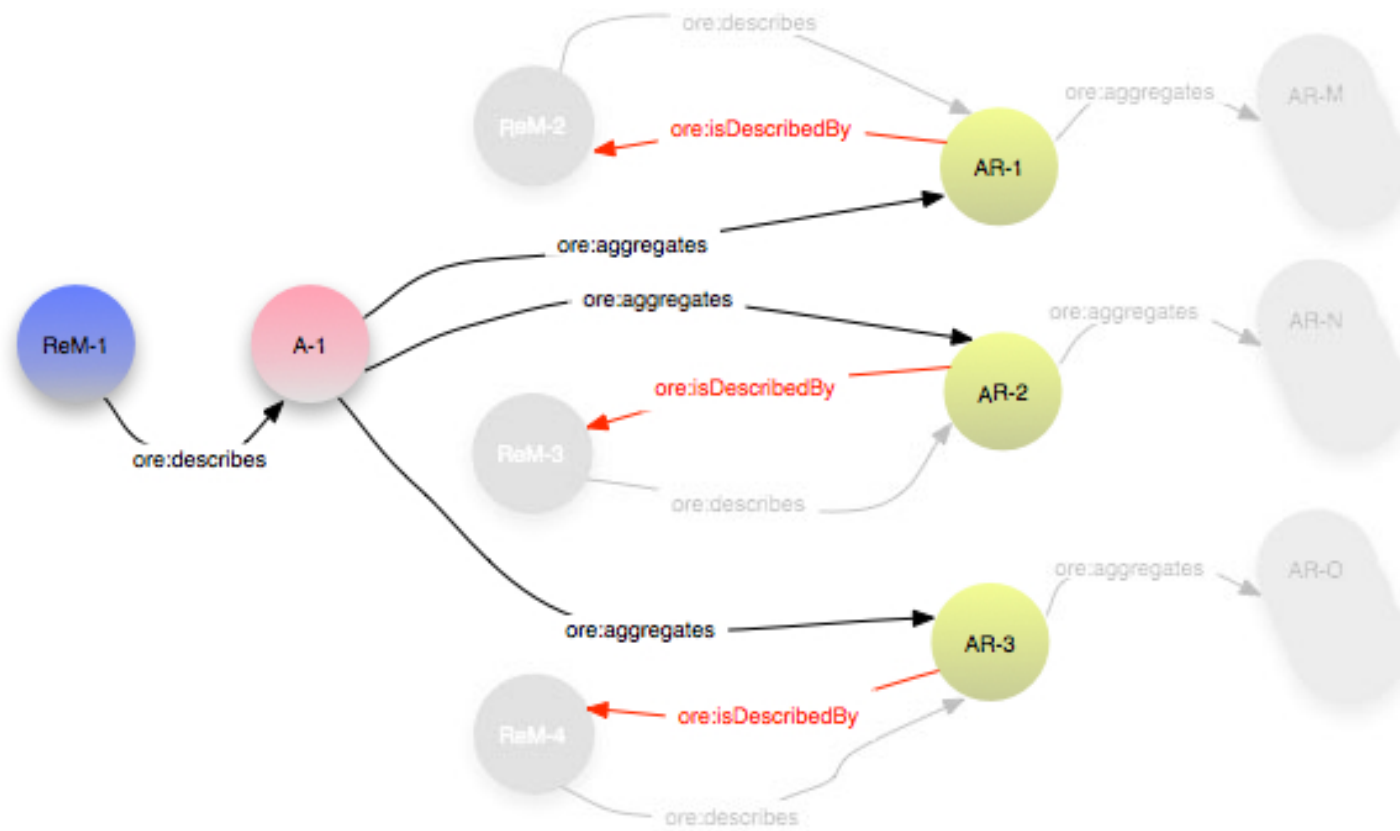
The ORE Data Model includes only the notion of a single level of aggregation. How do I represent aggregations of aggregations?



Structure within an Aggregation



Nested Aggregations



<ReM-1>	<ore:describe>	<A-1>
<A-1>	<ore:aggregates>	<AR-1>
<AR-1>	<ore:describedBy>	<ReM-2>
<A-1>	<ore:aggregates>	<AR-2>
<AR-2>	<ore:describedBy>	<ReM-3>
<A-1>	<ore:aggregates>	<AR-3>
<AR-3>	<ore:describedBy>	<ReM-4>

ore:describedBy	http://www.openarchives.org/ore/terms/describedBy
-----------------	---

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?

“It is clear that RDF/XML has already too many options in the ways to encode RDF graphs” [Dave Beckett, editor of RDF/XML specification

<http://www.w3.org/TR/rdf-syntax-grammar/> (2004) commenting on the revision process in

<http://www.dajobe.org/2003/05/iswc/paper.html> (ISWC 2003)]

Atom vs RDF Syntaxes

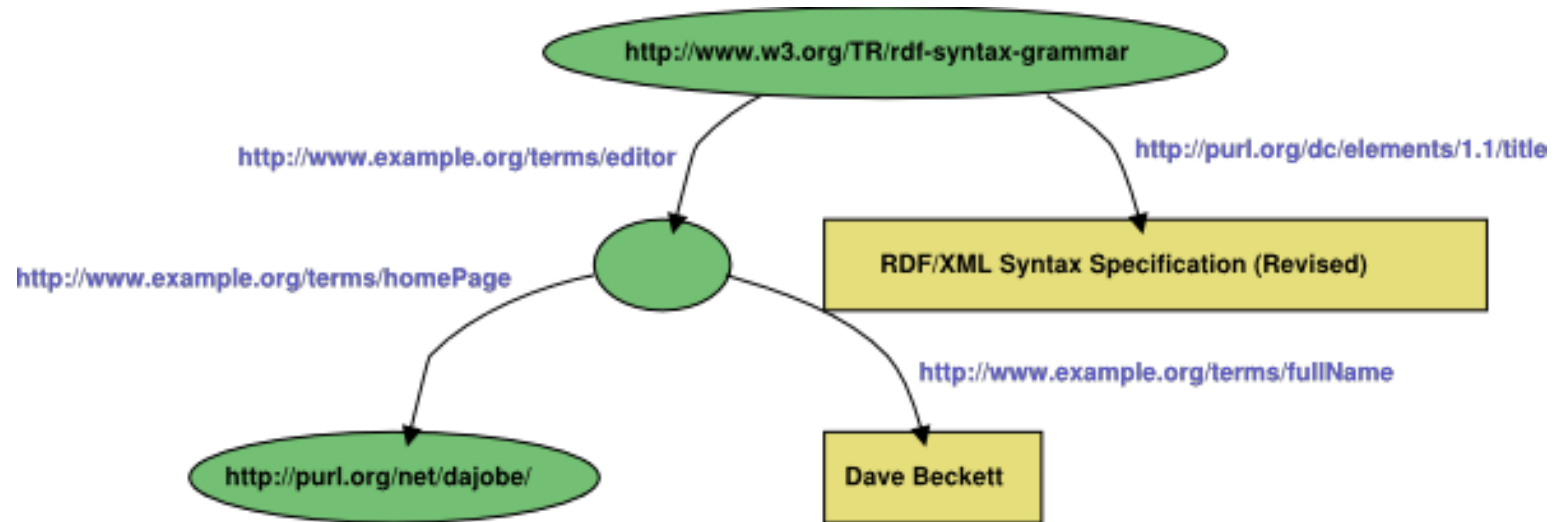
pro Atom

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

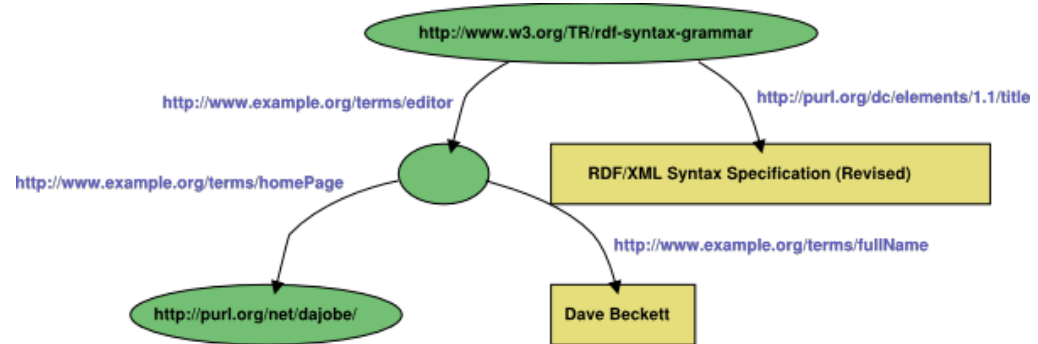
pro RDF Syntaxes

- Complete expressive power in one style
- No mapping required
- Easy extensibility
- RDFa can be embedded in XHTML

RDF/XML structure and serialization



RDF/XML nested or striped XML

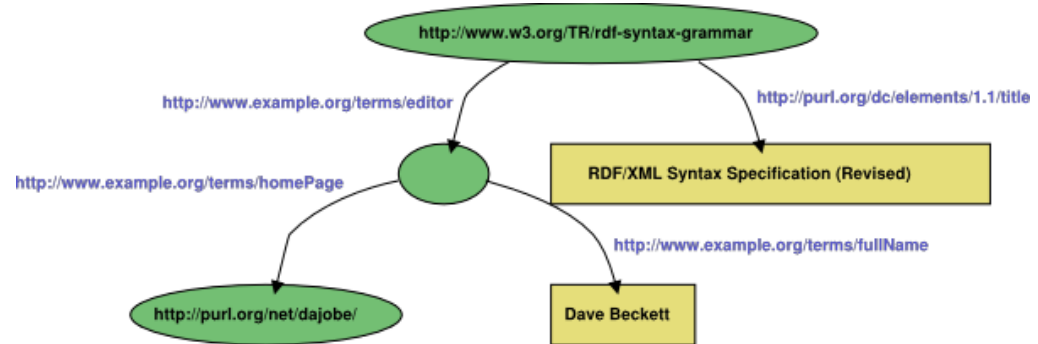


...

```
<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

Resource Map RDF/XML example

```
...  
    xml:base="http://dlib.org/dlib/february06/smith/02smith/"
```

```
...  
<rdf:Description rdf:about="rem">  
  <ore:describes rdf:resource="agg" />  
  <rdf:type rdf:resource="http://www.openarchives.org/ore/terms/ResourceMap" />  
  <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:type rdf:resource="http://purl.org/dc/dcmitype/Text" />  
</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
 - ⇒ need to be able to find Resource Map from Aggregation
 - ⇒ RULE: must be a mechanism to do this

Cool URIs

Aggregation	A-1	http://example.org/foo
Resource Map	ReM-1	http://example.org/foo.xml

How do we get from Aggregation to Resource Map?

Content Negotiation — ReM-1 URI in Content-Location

Redirection — 303 *ala* Linked Data

Aggregation	A-1	http://example.org/foo
Resource Map	ReM-1	http://example.org/foo.xml (Atom)
	ReM-2	http://example.org/foo.rdf (RDF/XML)

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.

Aggregation	A-1	http://example.org/foo.xml
Resource Map	ReM-1	http://example.org/foo.xml#rem

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

RDFa

Aggregation	A-1	http://example.org/foo
Resource Map	ReM-1	http://example.org/foo.html (+RDFa)
	ReM-2	http://example.org/foo.xml (Atom)
	ReM-3	http://example.org/foo.rdf (RDF/XML)

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)

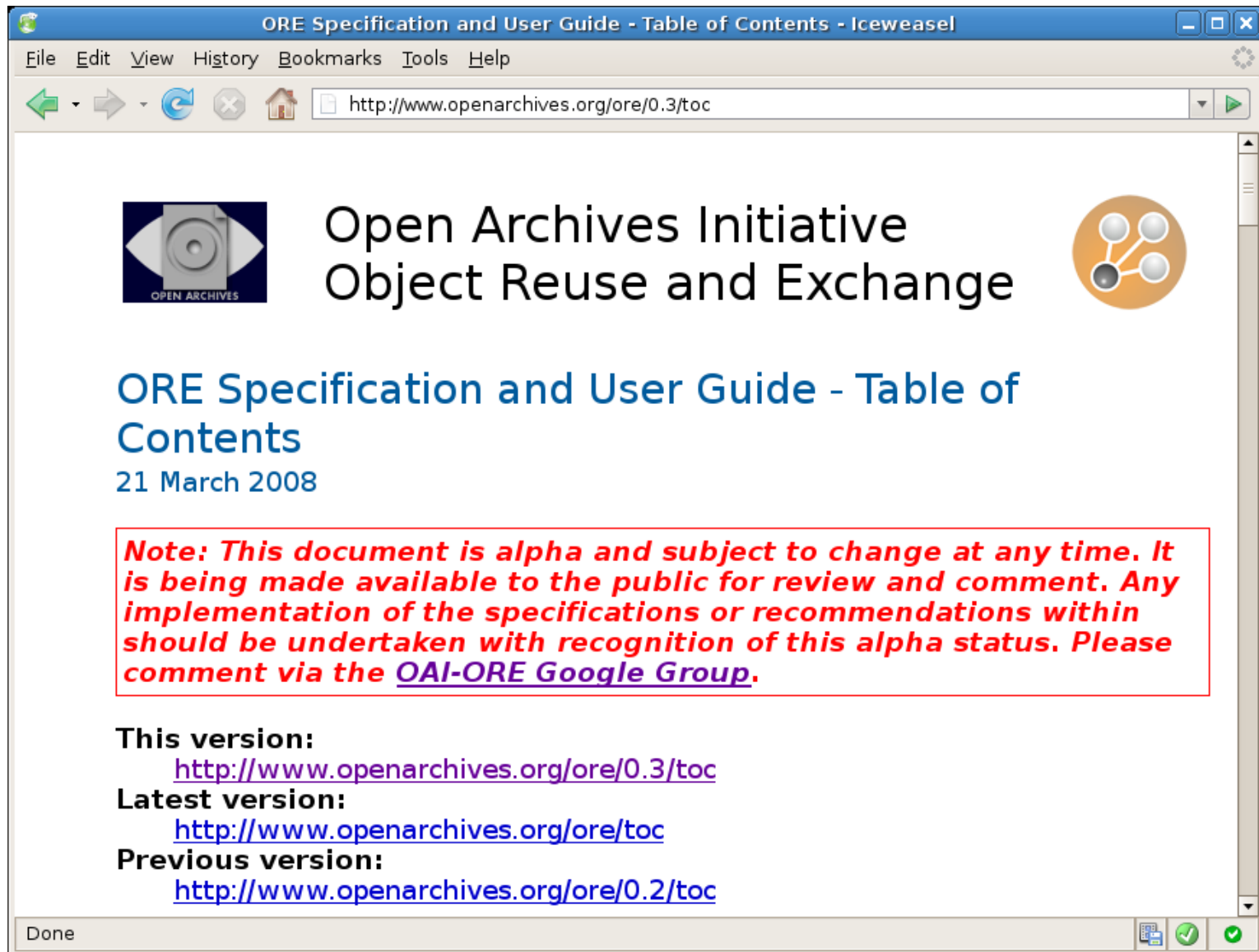
- **Validator for ORE Resource Maps in Atom** — alpha version available at

<http://www.openarchives.org/ore/atom-validator>

More tools

- **Schematron Schema for the Resource Map Profile** —
available at <http://www.openarchives.org/ore/atom-tron>
- **GRDDL crosswalk from Atom XML to RDF/XML** —
available at <http://www.openarchives.org/ore/atom-grddl>

7. Where to start...




The screenshot shows a web browser window with the title "ORE Specification and User Guide - Table of Contents - Iceweasel". The address bar contains the URL "http://www.openarchives.org/ore/0.3/toc". The page content includes the Open Archives Initiative logo, the text "Open Archives Initiative Object Reuse and Exchange", and the title "ORE Specification and User Guide - Table of Contents" dated "21 March 2008". A red-bordered box contains a note about the document's alpha status. Below this, there are links for "This version", "Latest version", and "Previous version".


ORE Specification and User Guide - Table of Contents - Iceweasel

File Edit View History Bookmarks Tools Help

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



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>

Done

... (and where I end)

- ORE User Guide - Primer for summary.
- ORE User Guide - Resource Map Implementation in Atom for most of what you need to know to create Resource Maps.
- ORE Specification - Resource Map Profile of Atom and ORE Specification - Abstract Data Model for the gory details. Use as reference.
- 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...

