Open Archives Initiative
Object Reuse & Exchange

# Tutorial
# Joint Conference on Digital Libraries 2008

Simeon Warner, Carl Lagoze,
Michael Nelson, Herbert Van de Sompel

Acknowledgments:
Pete Johnston, Rob Sanderson (co-editors of ORE documents)
Includes slides for Michael Kurtz, Astrophysics Data Service




# Tutorial Agenda

| | |
|---|---|
| 1:00pm | Context and Motivation (Simeon) |
| | Abstract Data Model and Serialization (Carl) |
| 2:30pm | Break, chat |
| 3:00pm | HTTP Implementation and Resource Map Discovery (Simeon) |
| | Building a Real Example (Carl) |
| 4:30pm | Continue with rest of life |




# Who we are

- Simeon Warner
  - simeon@cs.cornell.edu
  - http://www.cs.cornell.edu/people/simeon/
- Carl Lagoze
  - lagoze@cs.cornell.edu
  - http://www.cs.cornell.edu/lagoze





# These slides

http://openarchives.org/ore/Presentations/ORE-Tutorial-JCDL-main.pdf




# OAI-ORE: Support

- The Andrew W. Mellon Foundation
- The Coalition for Networked Information
- Joint Information Systems Committee
- Microsoft Corporation
- The National Science Foundation




# OAI -ORE: Technical Experts

**ORE Technical Committee**

| | |
|---|---|
| Chris Bizer | Freie Universität Berlin |
| Les Carr | University of Southampton |
| Tim DiLauro | Johns Hopkins University |
| Leigh Dodds | Ingenta |
| David Fulker | UCAR |
| Tony Hammond | Nature Publishing Group |
| Pete Johnston | Eduserv Foundation |
| Richard Jones | HP Labs |
| Carl Lagoze | Cornell University |
| Peter Murray | OhioLINK |
| Michael Nelson | Old Dominion University |
| Ray Plante | NCSA and National Virtual Observatory |
| Rob Sanderson | University of Liverpool |
| Herbert Van de Sompel | Los Alamos National Laboratory |
| Simeon Warner | Cornell University |
| Jeff Young | OCLC |

**ORE Liaison Group**

| | |
|---|---|
| Leonardo Candela | Consiglio Nazionale delle Ricerche - DRIVER |
| Tim Cole | University of Illinois Urbana-Champaign - Aquifer |
| Julie Allinson | JISC |
| Jane Hunter | University of Queensland - DEST |
| Savas Parastatidis | Microsoft Corporation |
| Sandy Payette | Fedora Commons |
| Thomas Place | University of Tilburg - DARE |
| Andy Powell | Eduserv Foundation - DCMI |
| Robert Tansley | Google, Inc. - DSpace |



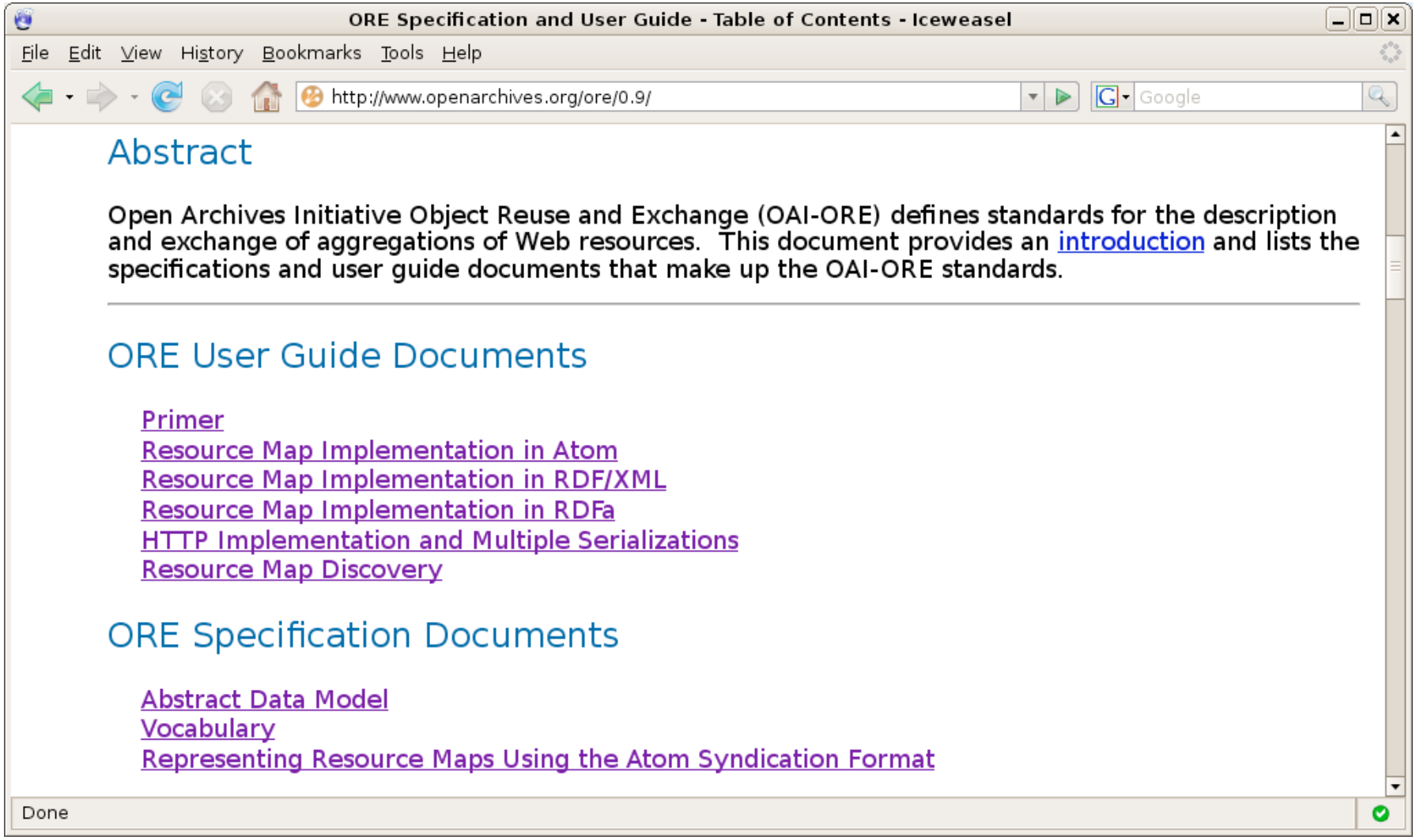

# Object Reuse and Exchange: Timeline

- Deliverables: http://www.openarchives.org/ore/toc
  - ORE Specifications alpha 0.1 (12/2007)
  - ORE Specifications alpha 0.2 (03/2008)
  - ORE Specifications alpha 0.3 (04/2008)
  - ORE Specifications beta (02/06/2008)
  - ORE Specification 1.0 (09/2008)

- Experiments to obtain feedback for specifications
  - 02/2008-08/2008

- Meetings:
  - March 3rd 2008, John Hopkins University: USA ORE Open Meeting
  - April 4th 2008, University of Southampton: European ORE Open Meeting
  - June 16th 2008, Tutorial at JCDL2008, Pittsburgh

# OAI-ORE: Documents



# Open Archives Initiative Object Reuse & Exchange

## Context and Motivation



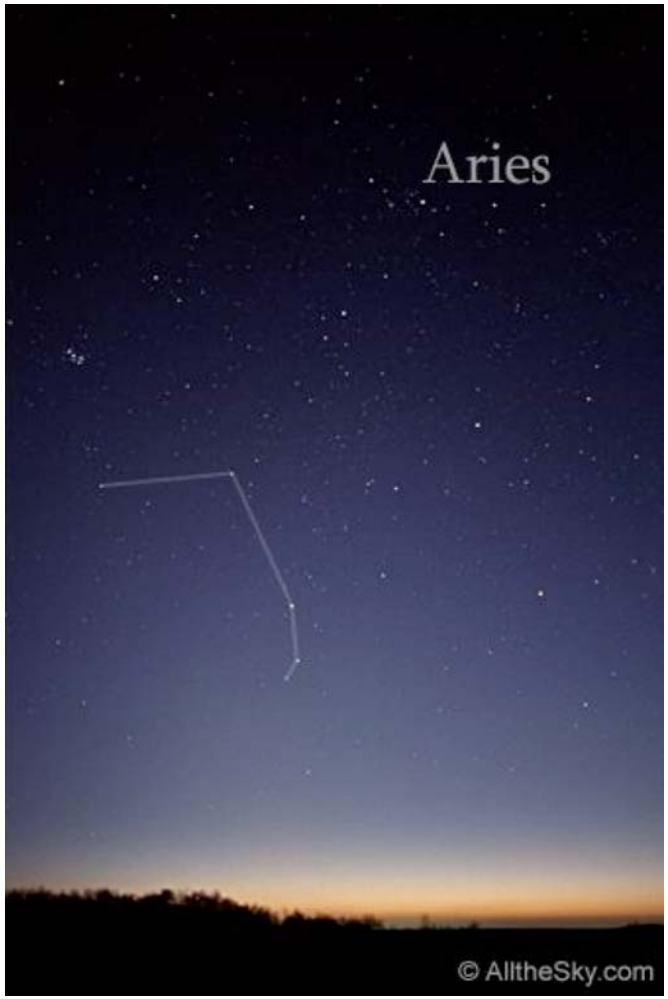

# OAI Object Reuse and Exchange

Subject: Aggregations of Web resources

Approach: Publish Resource Maps to the Web that Instantiate, Describe, and Identify Aggregations








# Instantiate, Describe, and Identify Aggregations

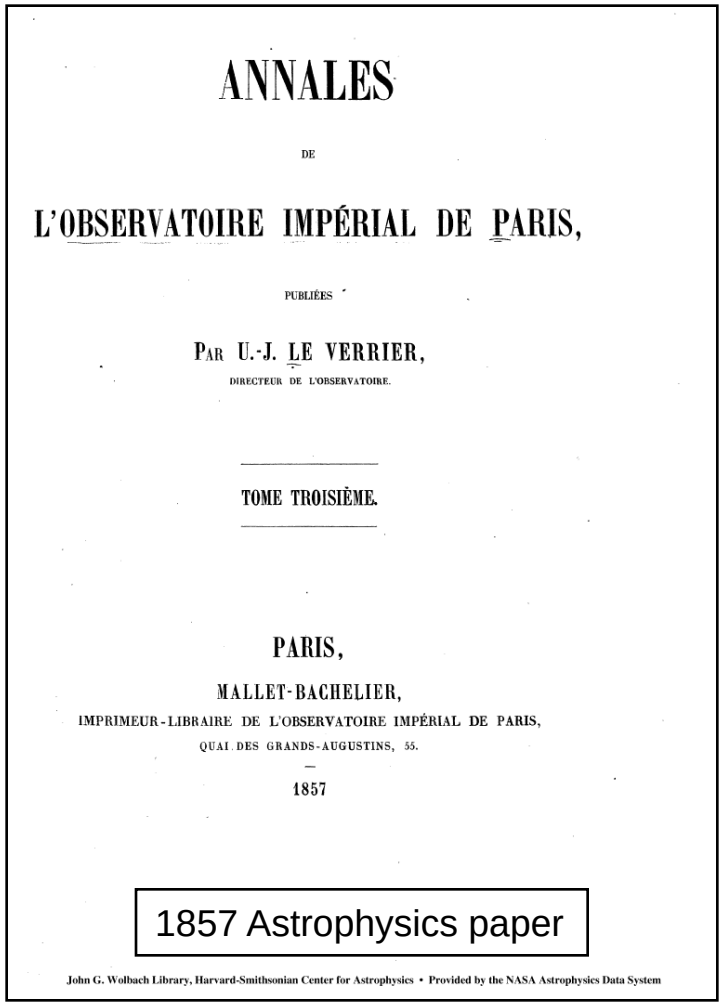
# Aggregations

It used to be that all information that was to be conveyed could be provided in a single container.

Babylonian Astronomical Catalogue

# Aggregations

It used to be that all information that was to be conveyed could be provided in a single container.



ANNALES

DE

L'OBSERVATOIRE IMPÉRIAL DE PARIS,

PUBLIÉES

PAR U.-J. LE VERRIER,

DIRECTEUR DE L'OBSERVATOIRE.

TOME TROISIÈME.

PARIS,

MALLET-BACHELIER,

IMPRIMEUR-LIBRAIRE DE L'OBSERVATOIRE IMPÉRIAL DE PARIS,

QUAI DES GRANDS-AUGUSTINS, 55.

1857

1857 Astrophysics paper



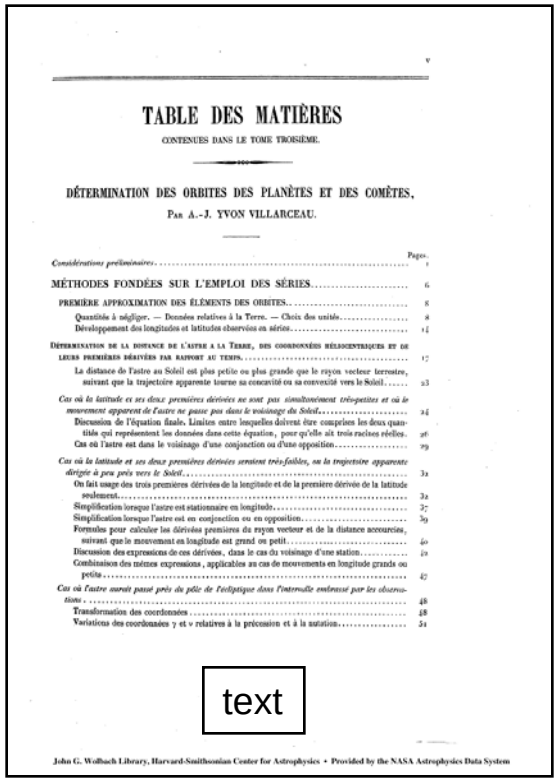


TABLE DES MATIÈRES

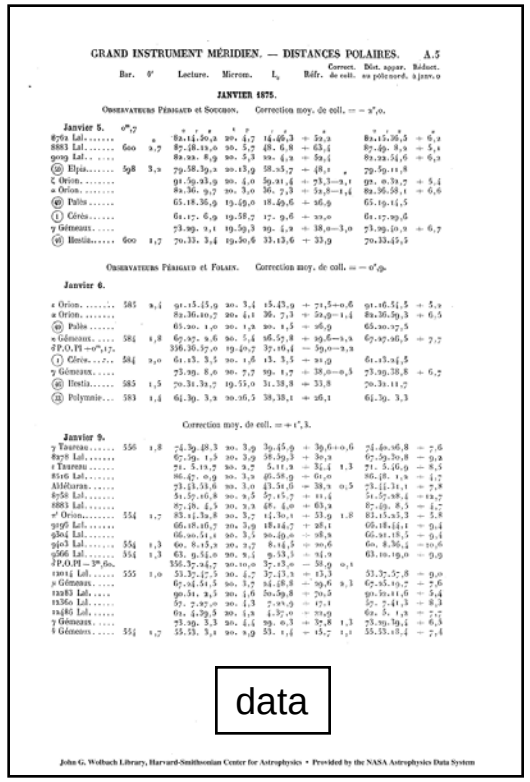
CONTENUES DANS LE TOME TROISIÈME.

DÉTERMINATION DES ORBITES DES PLANÈTES ET DES COMÈTES,

PAR A.-J. YVON VILLARCEAU.

text





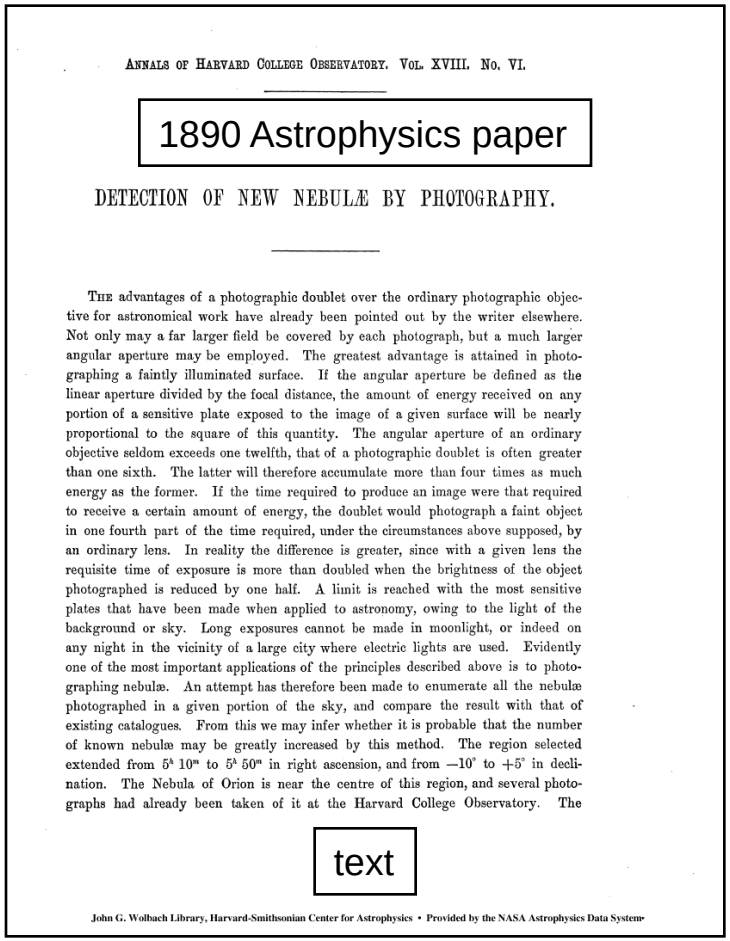
GRAND INSTRUMENT MÉRIDIEN. — DISTANCES POLAIRES.

JANVIER 1875.

data

# Aggregations

In scholarly communication that didn't last very long.



ANNALS OF HARVARD COLLEGE OBSERVATORY. VOL. XVIII. NO. VI.

1890 Astrophysics paper

DETECTION OF NEW NEBULÆ BY PHOTOGRAPHY.

THE advantages of a photographic doublet over the ordinary photographic objective for astronomical work have already been pointed out by the writer elsewhere. Not only may a far larger field be covered by each photograph, but a much larger angular aperture may be employed. The greatest advantage is attained in photographing a faintly illuminated surface. If the angular aperture be defined as the linear aperture divided by the focal distance, the amount of energy received on any portion of a sensitive plate exposed to the image of a given surface will be nearly proportional to the square of this quantity. The angular aperture of an ordinary objective seldom exceeds one twelfth, that of a photographic doublet is often greater than one sixth. The latter will therefore accumulate more than four times as much energy as the former. If the time required to produce an image were that required to receive a certain amount of energy, the doublet would photograph a faint object in one fourth part of the time required, under the circumstances above supposed, by an ordinary lens. In reality the difference is greater, since with a given lens the requisite time of exposure is more than doubled when the brightness of the object photographed is reduced by one half. A limit is reached with the most sensitive plates that have been made when applied to astronomy, owing to the light of the background or sky. Long exposures cannot be made in moonlight, or indeed on any night in the vicinity of a large city where electric lights are used. Evidently one of the most important applications of the principles described above is to photographing nebulæ. An attempt has therefore been made to enumerate all the nebulæ photographed in a given portion of the sky, and compare the result with that of existing catalogues. From this we may infer whether it is probable that the number of known nebulæ may be greatly increased by this method. The region selected extended from $5^h\ 10^m$ to $5^h\ 50^m$ in right ascension, and from $-10°$ to $+5°$ in declination. The Nebula of Orion is near the centre of this region, and several photographs had already been taken of it at the Harvard College Observatory. The

text



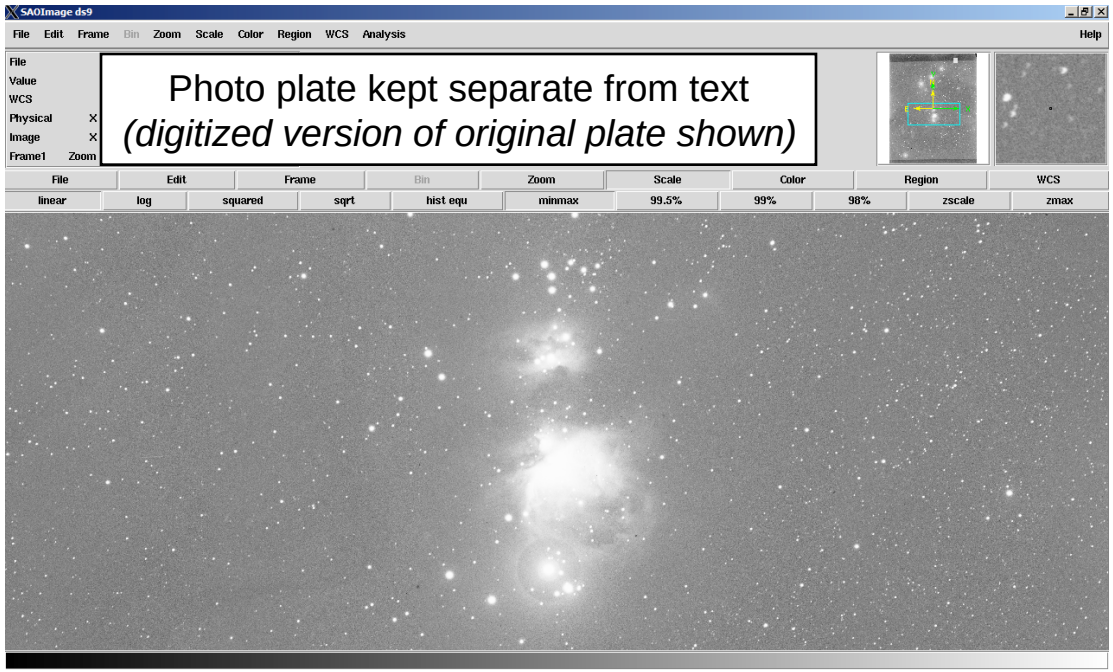



Photo plate kept separate from text
*(digitized version of original plate shown)*

# Aggregations

2006 Astrophysics paper

ENTROPY PROFILES IN THE CORES OF COOLING FLOW CLUSTERS OF GALAXIES

Megan Donahue,[1] Donald J. Horner,[2] Kenneth W. Cavagnolo,[1] and G. Mark Voit[1]

*Received 2005 July 13; accepted 2006 February 6*

ABSTRACT

The X-ray properties of a relaxed cluster of galaxies are determined primarily by its gravitational potential well and the entropy distribution of its intracluster gas. That entropy distribution reflects both the accretion history of the cluster and the feedback processes that limit the condensation of intracluster gas. Here we present *Chandra* observations of the core entropy profiles of nine classic "cooling flow" clusters that appear relatively relaxed (at least outside the central 10–20 kpc) and contain intracluster gas with a cooling time less than a Hubble time. We show that those entropy profiles are remarkably similar, despite the fact that the clusters range over a factor of 3 in temperature. They typically have an entropy level of $\approx$130 keV cm$^2$ at 100 kpc that declines to a plateau $\sim$10 keV cm$^2$ at $\leq$10 kpc. Between these radii, the entropy profiles are $\propto r^{\alpha}$ with $\alpha \approx 1.0$–1.3. The nonzero central entropy levels in these clusters correspond to a cooling time $\sim 10^8$ yr, suggesting that episodic heating on this timescale maintains the central entropy profile in a quasi–steady state. We show in an appendix that although disturbances and bubbles are visible in the central regions of these clusters, these phenomena do not strongly bias our entropy estimates.

*Subject headings:* catalogs — cosmology: observations — galaxies: clusters: general — methods: data analysis — X-rays: galaxies: clusters

*Online material:* color figures

1. INTRODUCTION

The global properties of a cluster of galaxies, such as its bolometric X-ray luminosity $L_X$ and its mean temperature $T_X$, are determined primarily by the mass $M_{vir}$ within a suitably chosen virial radius. A cluster's temperature depends on mass because mass determines the depth of the cluster's potential well. Its X-ray luminosity depends on mass because mass determines both the total number of baryons in the cluster and the potential well confining those baryons. However, several secondary factors combine to produce a dispersion in both $L_X$ and $T_X$ at a fixed $M_{vir}$, and understanding the nature of that dispersion is crucial to doing precision cosmology with clusters. One of those factors is merger shocks, which can temporarily raise both the luminosity and best-fitting temperature of a cluster (e.g., Randall et al. 2002). A second is the shape of the potential well, because clusters whose potentials are more centrally concentrated tend to have higher central temperatures (e.g., Voit et al. 2002). A third factor is the amount of intracluster gas with a cooling time less than the age of the universe. The presence of such gas leads to both a large peak in the central surface brightness of a cluster and a central temperature gradient that rises with radius. Consequently, clusters having larger amounts of gas with a short cooling time tend to have higher $L_X$ and lower $T_X$ at a given value of $M_{vir}$ (Allen & Fabian 1998; Fabian et al. 1994; Markevitch 1998).

Such clusters have often been called "cooling flow clusters" because the central gas was thought to condense and flow toward the center of the cluster as it radiated away its thermal energy (for a recent review see Donahue & Voit 2004). Observations from *Chandra* and *XMM-Newton* now show that the central gas is not simply cooling to low temperatures and condens

manner originally envisioned (e.g., Peterson et al. 2001, 2003). Some form of feedback apparently prevents the central gas from condensing and forming stars, thereby truncating the high end of the galaxy luminosity function. The nature of that feedback is currently an active topic of both observational and theoretical research, focusing largely on the role of outflows from active galactic nuclei (AGNs) in cluster cores.

This paper analyzes archival *Chandra* data on nine cooling flow clusters seeking clues to what keeps that gas from condensing and why clusters of a given mass have different amounts of gas with a short central cooling time. The tactic we take in our analysis is to focus on the entropy profiles of these clusters. We concentrate on entropy because it is a more fundamental property of the intracluster medium (ICM) itself than either temperature or density alone. For example, the temperature of a cluster's gas primarily reflects the cluster's potential well depth; heating or cooling of the gas merely causes it to expand or contract in the potential well with only a modest change in temperature. The density of that gas depends on how much gravity can compress it in the cluster's potential well, and it is the specific entropy of the gas that determines its density at a given pressure. Thus, the observable X-ray properties of a relaxed cluster of galaxies depend almost entirely on two physical attributes: (1) the shape and depth of the cluster's dark matter halo and (2) the entropy distribution of the intracluster gas (e.g., Voit et al. 2002).

Intracluster entropy is also intimately related to the cooling and feedback processes that govern galaxy evolution and that may also play a role in limiting condensation in cluster cores. Theories and simulations of cluster formation that ignore these processes fail to reproduce the observable properties of present-day clusters. ity alone were responsible for shaping the appearances of and groups, then we would expect their properties to be self-similar, with a luminosity-temperature relation like . Furthermore, we would expect groups and clusters similar surface brightness profiles, when scaled to the virial radius of the system. However, observations indicate that

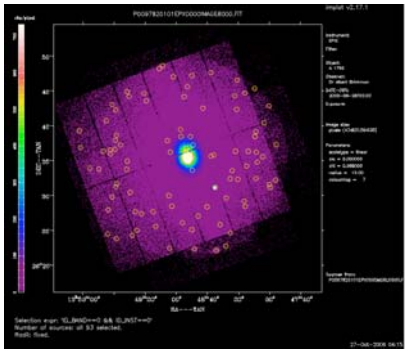
[1] Department of Physics and Astronomy, Michigan State Univ Building, East Lansing, MI 48824; donahue@pa.msu.edu, cavagno .edu, voit@pa.msu.edu.

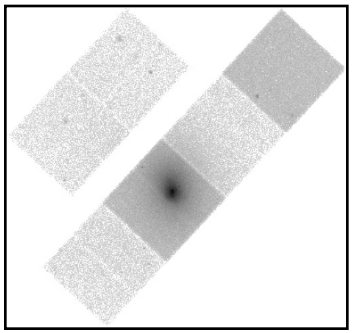
[2] NASA Goddard Space Flight Center, Code 660, Greenbelt, horner@milkyway.gsfc.nasa.gov.



text

And in digital scholarly communication, the single container concept is obsolete.





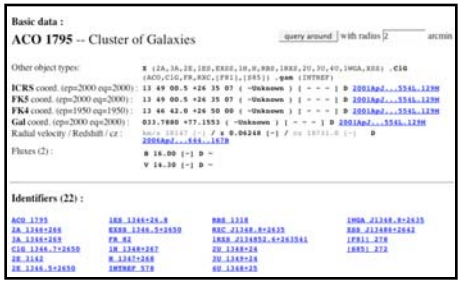
X-MM-Newton X-ray observation
Vilspa, Spain

A1795

Chandra X-ray observation
Cambridge, MA

Basic object information
Strasbourg, France

Hubble optical observation
Baltimore, MD



Basic data :

ACO 1795 -- Cluster of Galaxies

query around | with radius | arcmin

Other object types:

ICRS coord. (ep=2000 eq=2000) : 13 49 00.5 +26 35 07 ( ~Unknown ) [ ~ ~ ~ ] D 2001ApJ...554L.129M

FK5 coord. (ep=2000 eq=2000) : 13 49 00.5 +26 35 07 ( ~Unknown ) [ ~ ~ ~ ] D 2001ApJ...554L.129M

FK4 coord. (ep=1950 eq=1950) : 13 46 42.0 +26 50 00 ( ~Unknown ) [ ~ ~ ~ ] D 2001ApJ...554L.129M

Gal coord. (ep=2000 eq=2000) : 033.7880 +77.1553 ( ~Unknown ) [ ~ ~ ~ ] D 2001ApJ...554L.129M

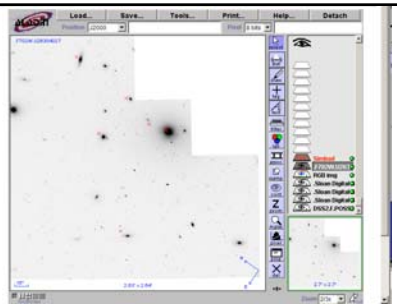
Radial velocity / Redshift / cz :

Fluxes (2) :

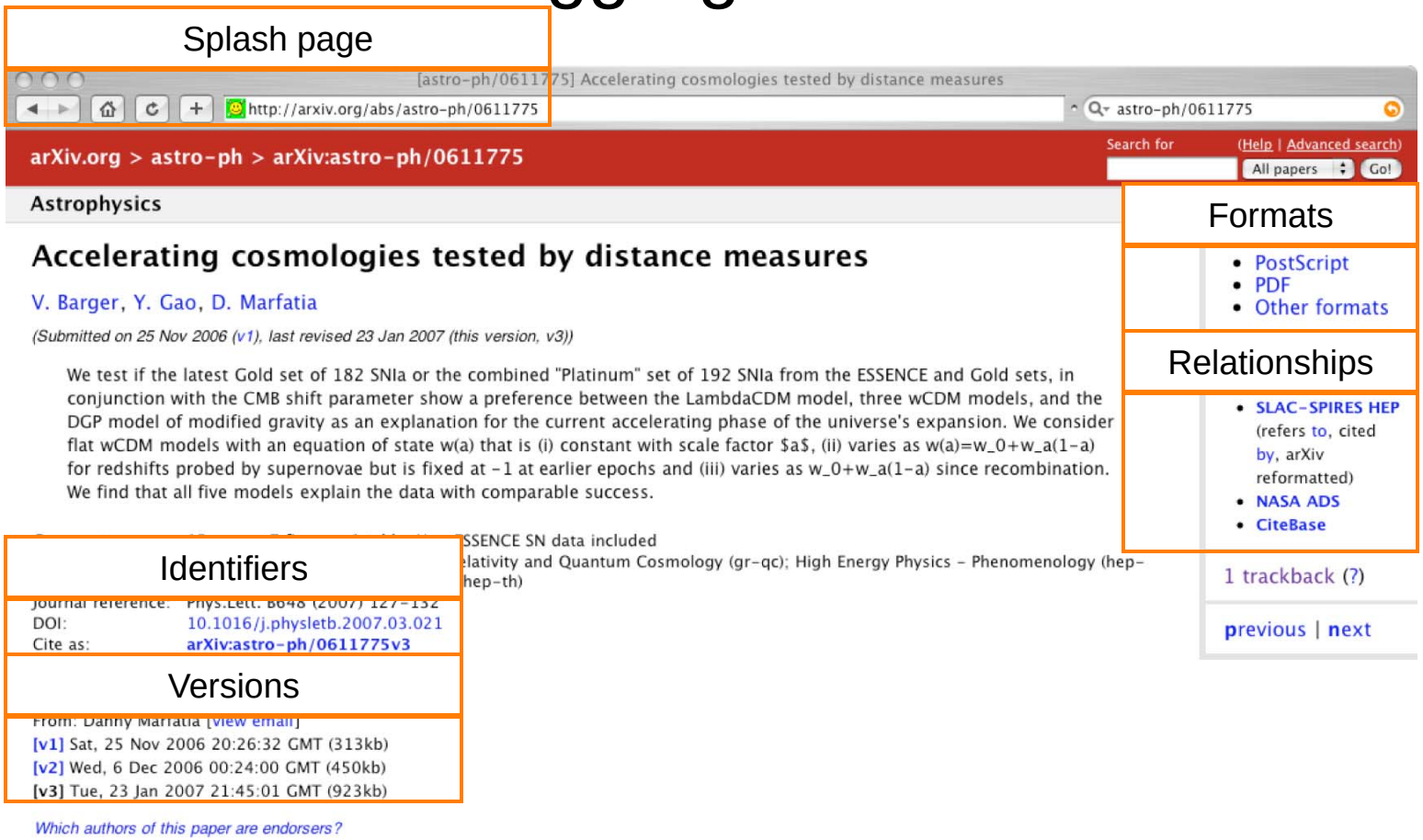
B 16.00 [~] D ~

V 14.30 [~] D ~

Identifiers (22) :

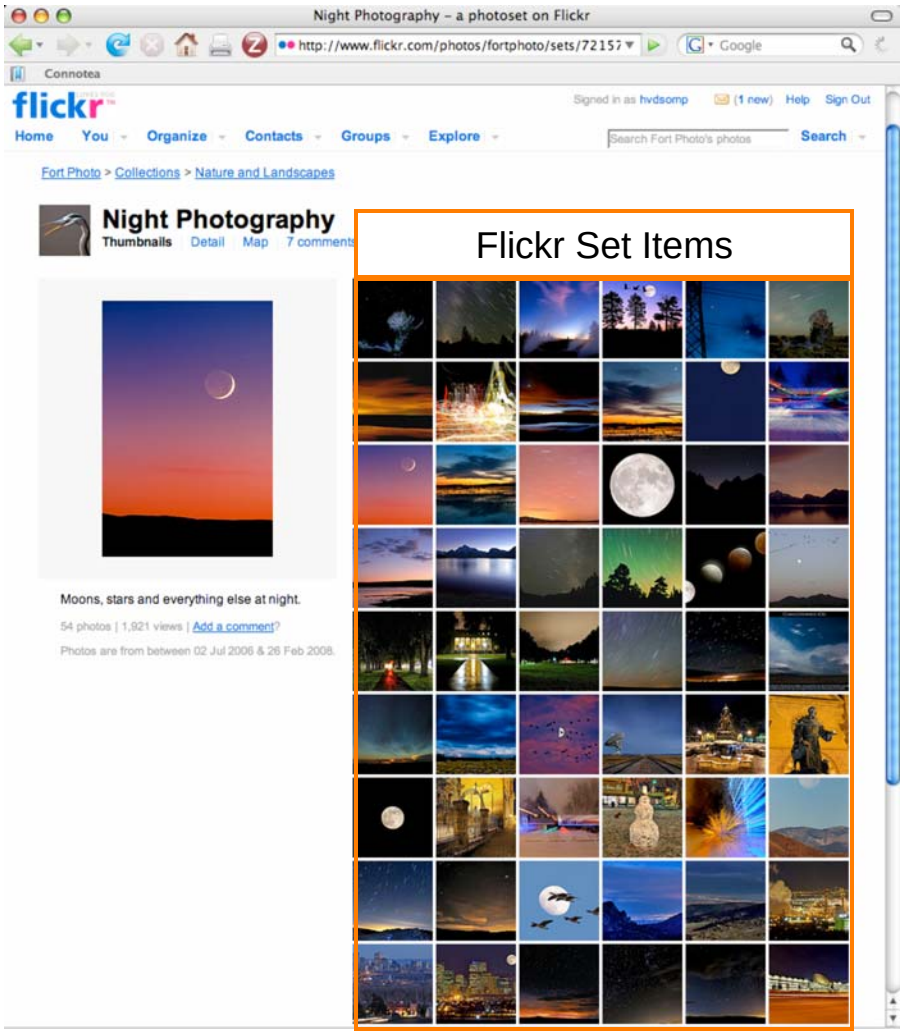


# Aggregations!



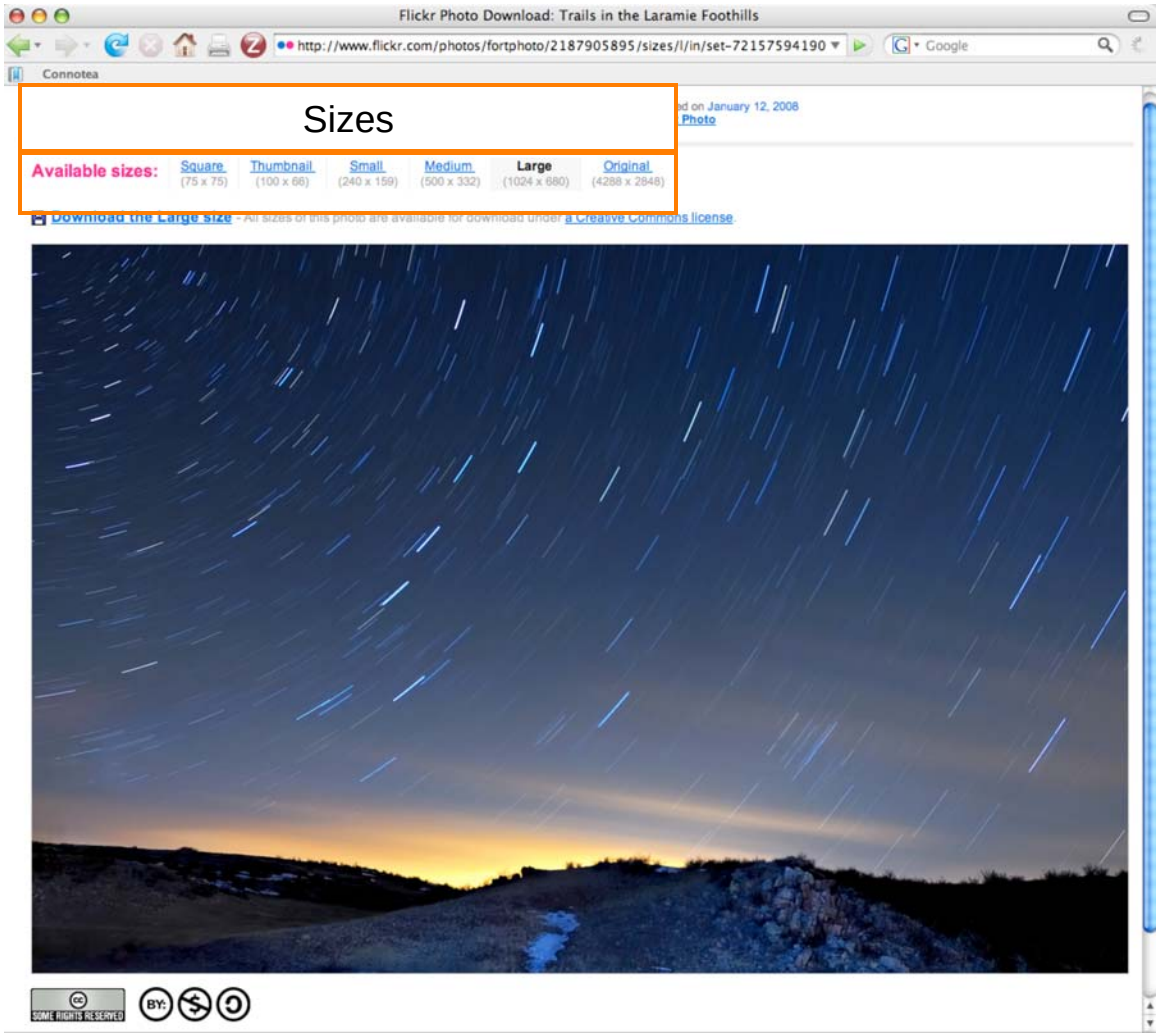

http://arxiv.org/abs/astro-ph/0611775

# Aggregations!!




http://www.flickr.com/photos/fortphoto/sets/72157594190371016/

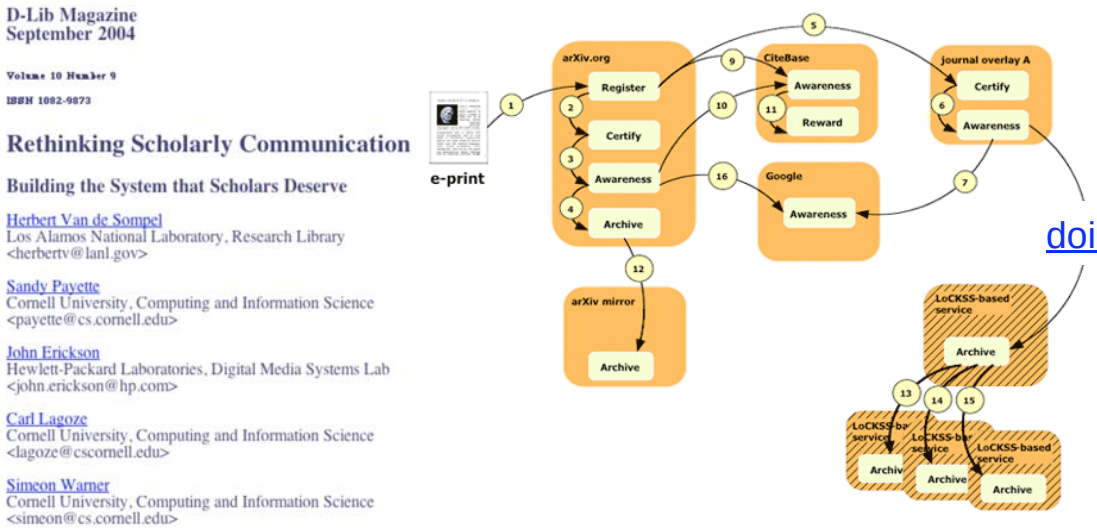
# Aggregations!!!




http://www.flickr.com/photos/fortphoto/sets/72157594190371016/

# Original Vision of Object Reuse and Exchange

- Scholarly communication as a global, cross-repository workflow.
  - Leverage the intrinsic value of the materials that become available in distributed repositories.
  - Value chains across repositories and applications with repository materials as their subject.
  - Make repositories **active nodes in a global environment**, not passive local nodes.
  - Life for those materials **starts** in repositories; it does not end there.
  - Materials from repositories must be **reusable in different contexts**.




doi:10.1045/september2004-vandesompel

# OAI-ORE: The Reality

Subject: Aggregations of Web resources

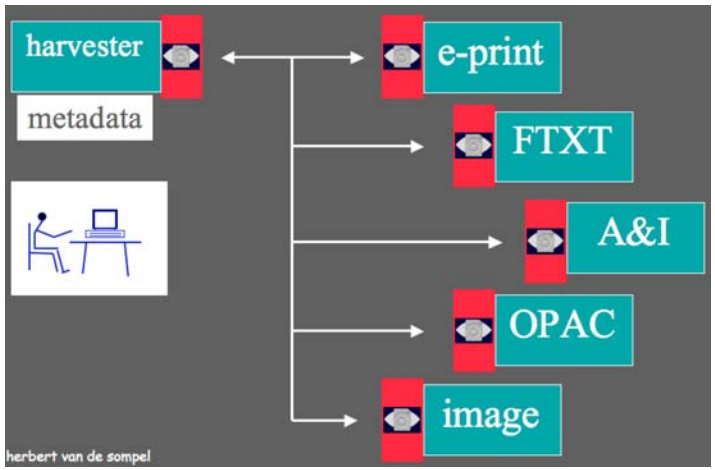
Approach: Publish Resource Maps to the Web that Instantiate, Describe, and Identify Aggregations

Reuse: URI of Aggregation as handle; Resource Map as the ore for value chains

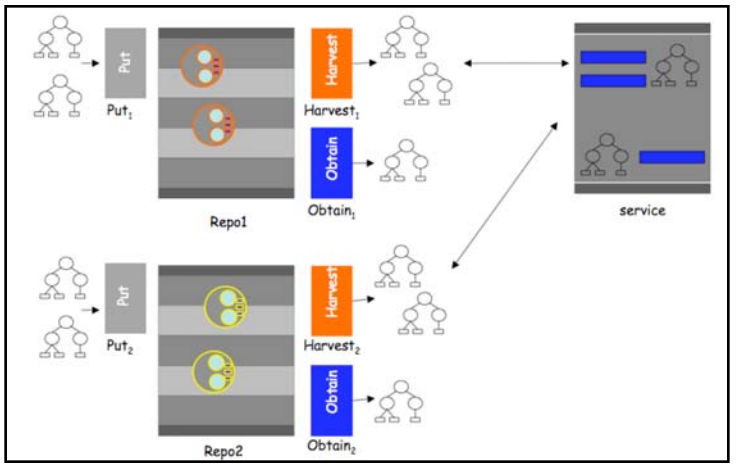



# OAI-ORE: A Resource-Centric Approach

- Prior efforts had the repository as the center of the interoperability thinking:
  - Including OAI-PMH
  - Including initial OAI-ORE thinking cf. “Augmenting Interoperability across Scholarly Repositories”
- This approach does not vibe well with the Web:
  - The Web Architecture knows resources and URIs, not repositories
  - Requires special treatment by applications that dominate the Web.

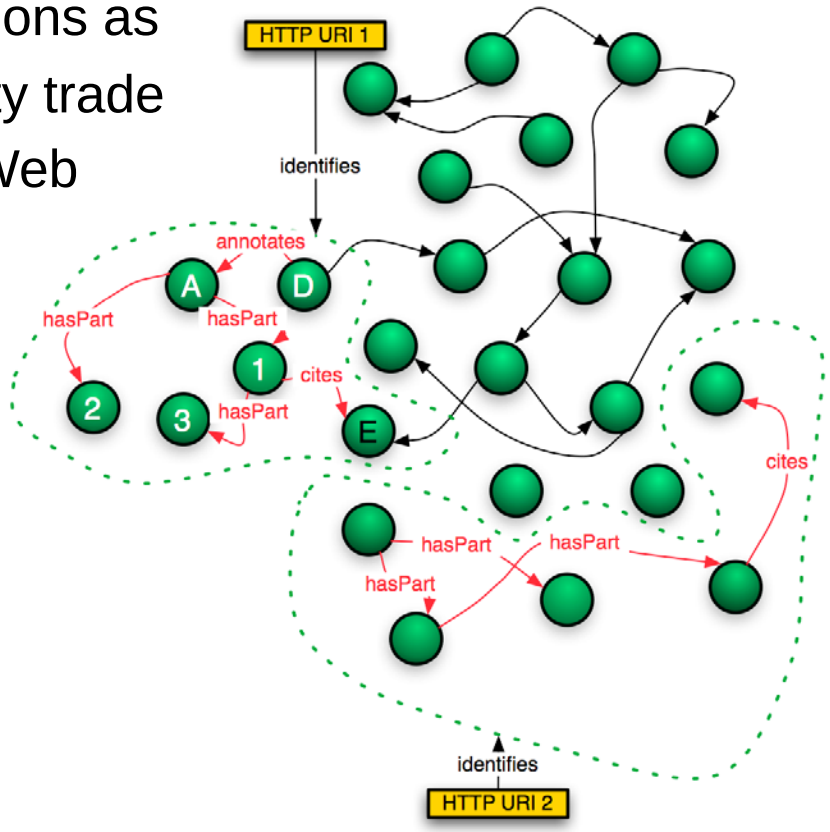



Keep dreaming!

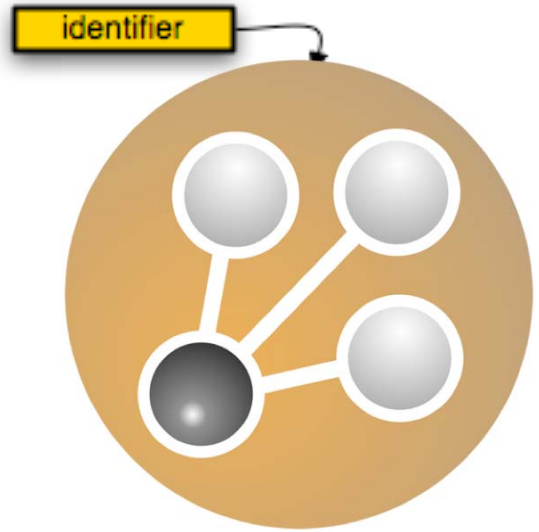


# OAI-ORE: A Resource-Centric Approach

- Fundamental shift in the chosen approach towards interoperability
- The Web Architecture as the platform for interoperability
- Resources, URIs, and representations as the tools of the ORE interoperability trade
- De-facto integration with existing Web applications
- Potential of adoption by other communities
- Potential of tools created by other communities
- ….



# From Compound Information Objects to Aggregations

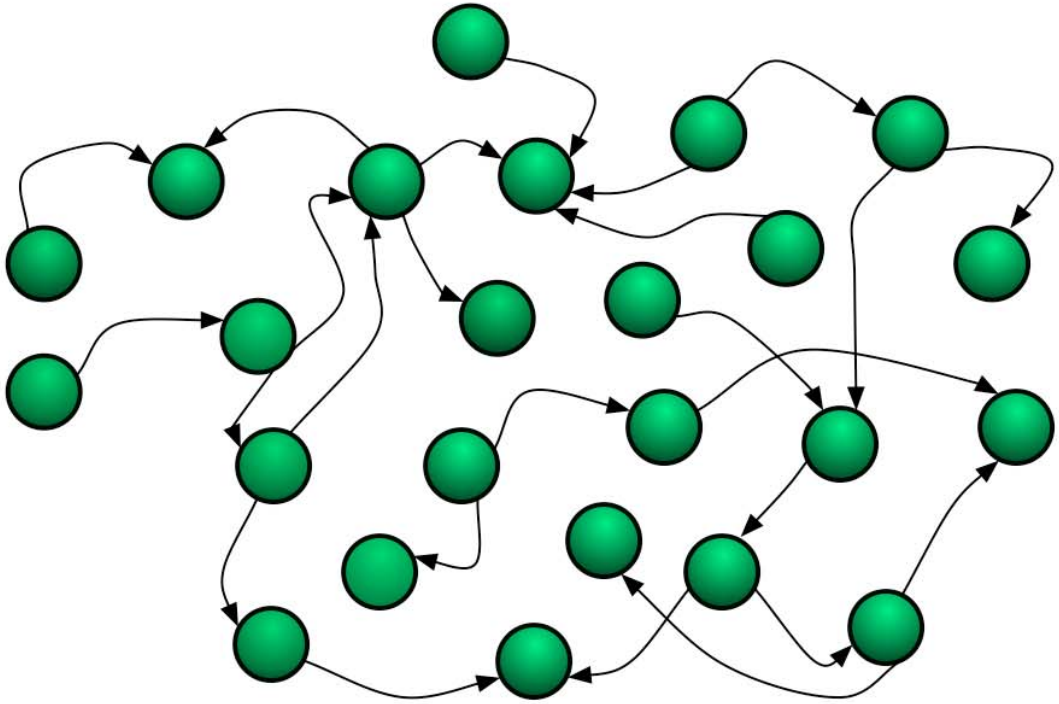



Identified, bounded aggregations of related information units that form a logical whole.

Components of a compound object may vary according to:

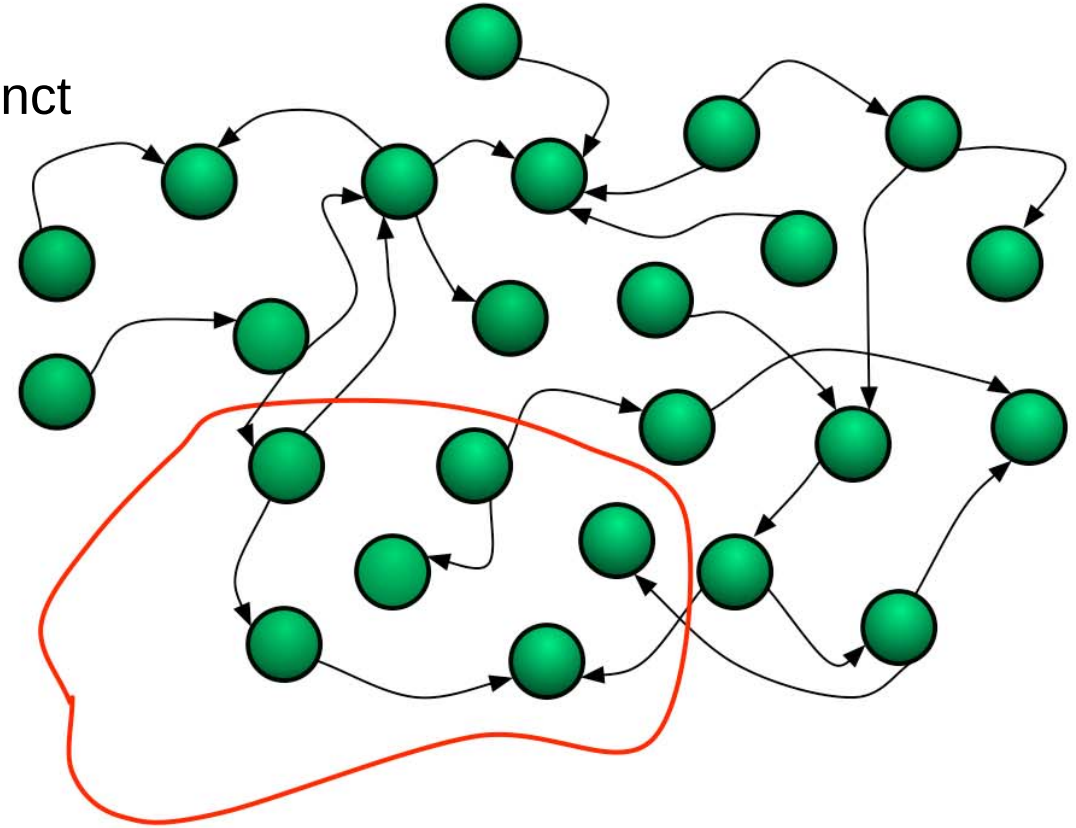
- Semantic type: book, article, software, dataset, simulation, …
- Media type: text, image, audio, video, mixed
- Media format: PDF, HTML, JPEG, MP3, …
- Network location
- Relationships: internal, external

# The Web



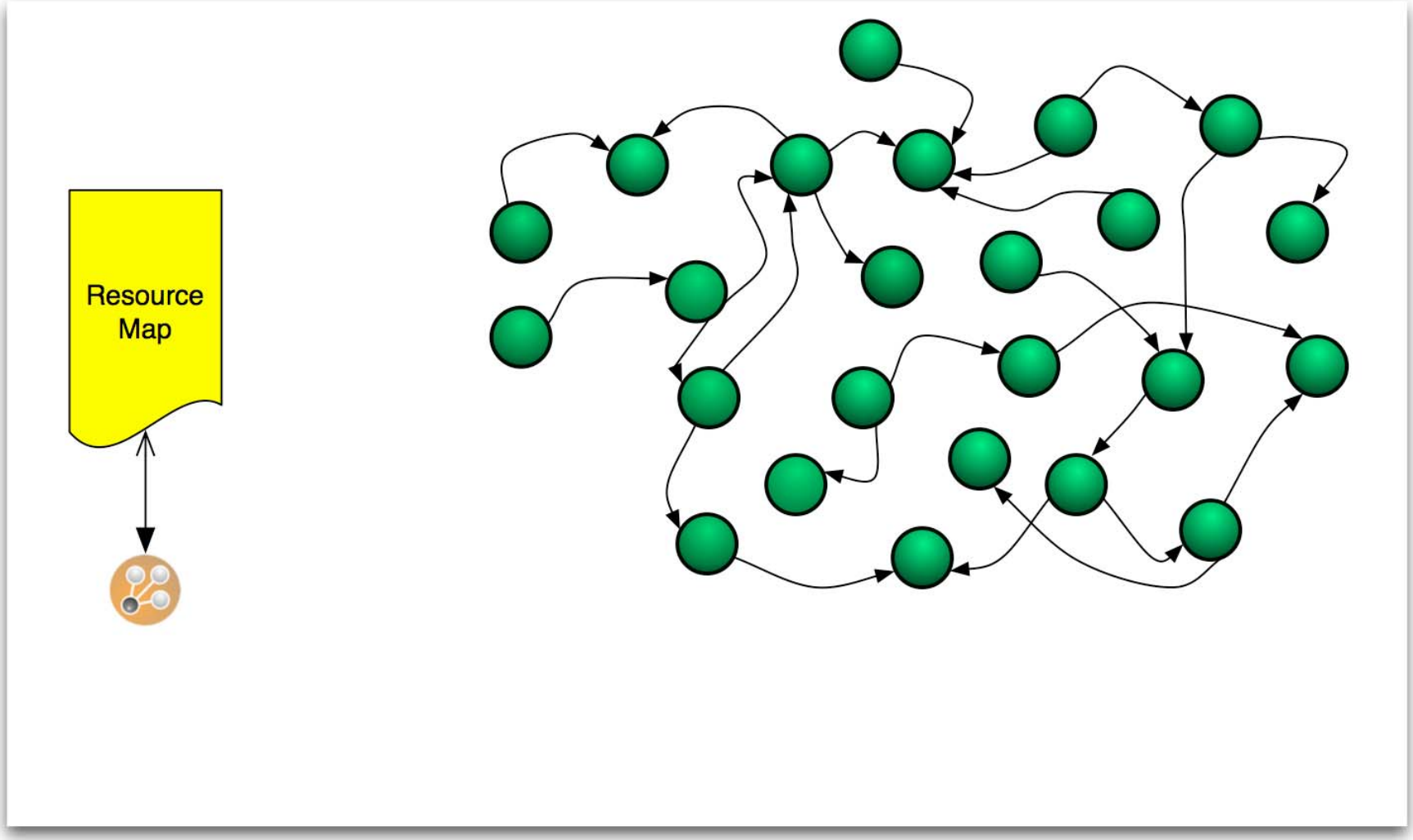
# An Aggregation and the Web

- Resources of an Aggregation are distinct URI-identified Web resources

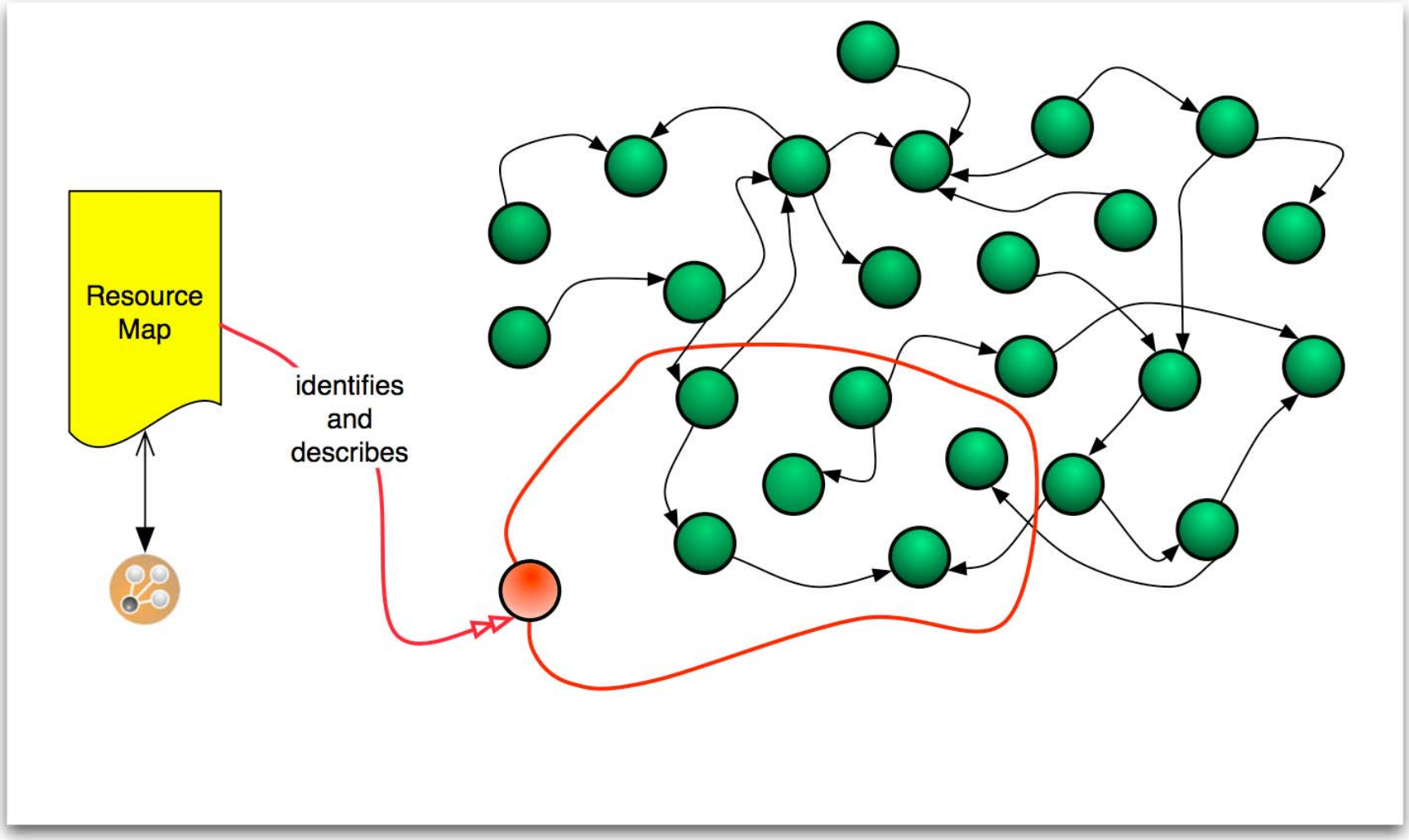


- Missing are:
  - The boundary that delineates the Aggregation in the Web
  - An identity (URI) for the Aggregation

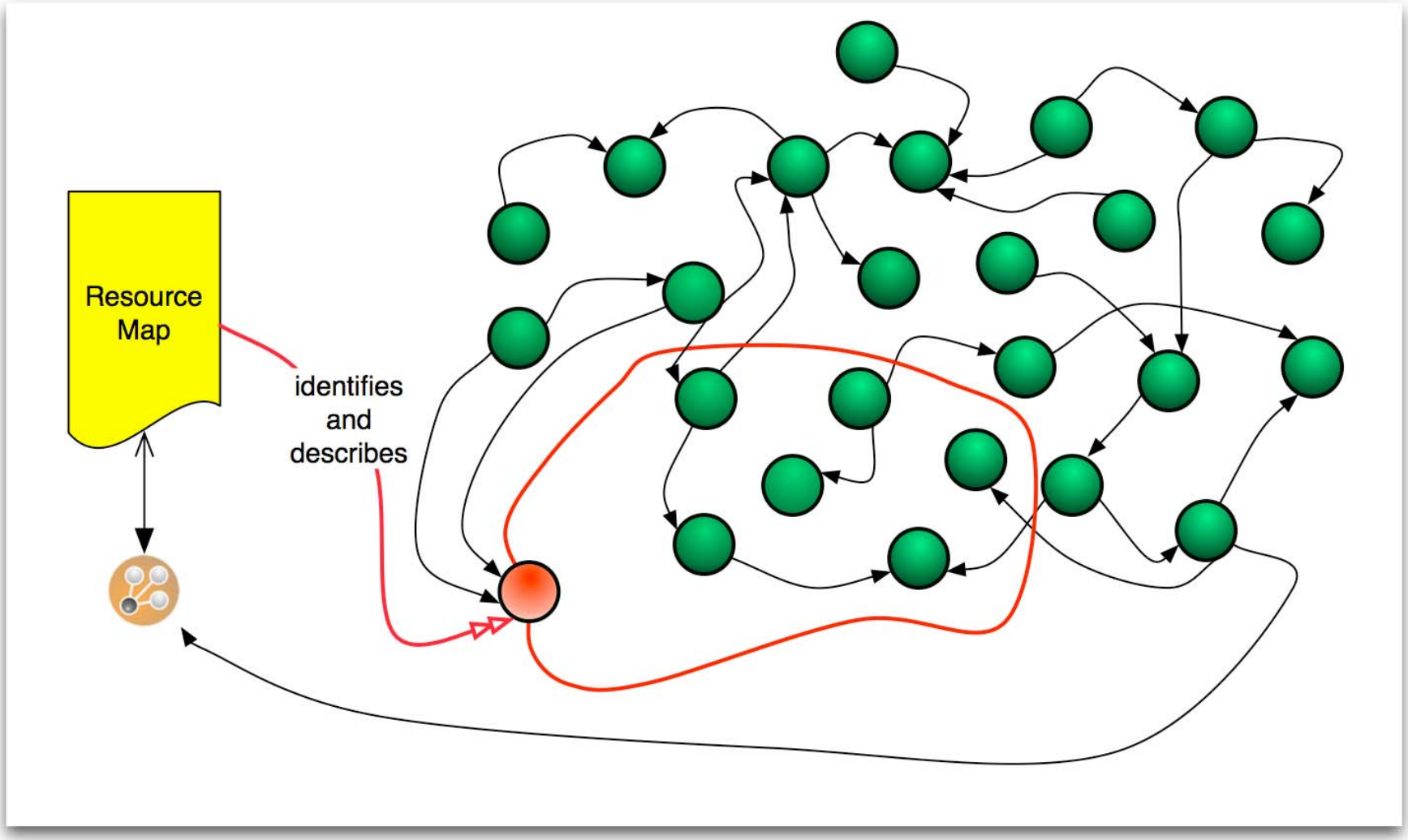
# Publish a Resource Map to the Web



# The Resource Map Identifies and Describes the Aggregation



# The Resource Map and the Aggregation integrate into the Web



# Approach and Agenda

Subject: Aggregations of Web resources

Approach: Publish Resource Maps to the Web that Instantiate, Describe, and Identify Aggregations

Reuse: URI of Aggregation as handle; Resource Map as the ore for value chains

Next... The ORE Data Model




# Open Archives Initiative Object Reuse & Exchange

## Basics:

## Abstract Data Model

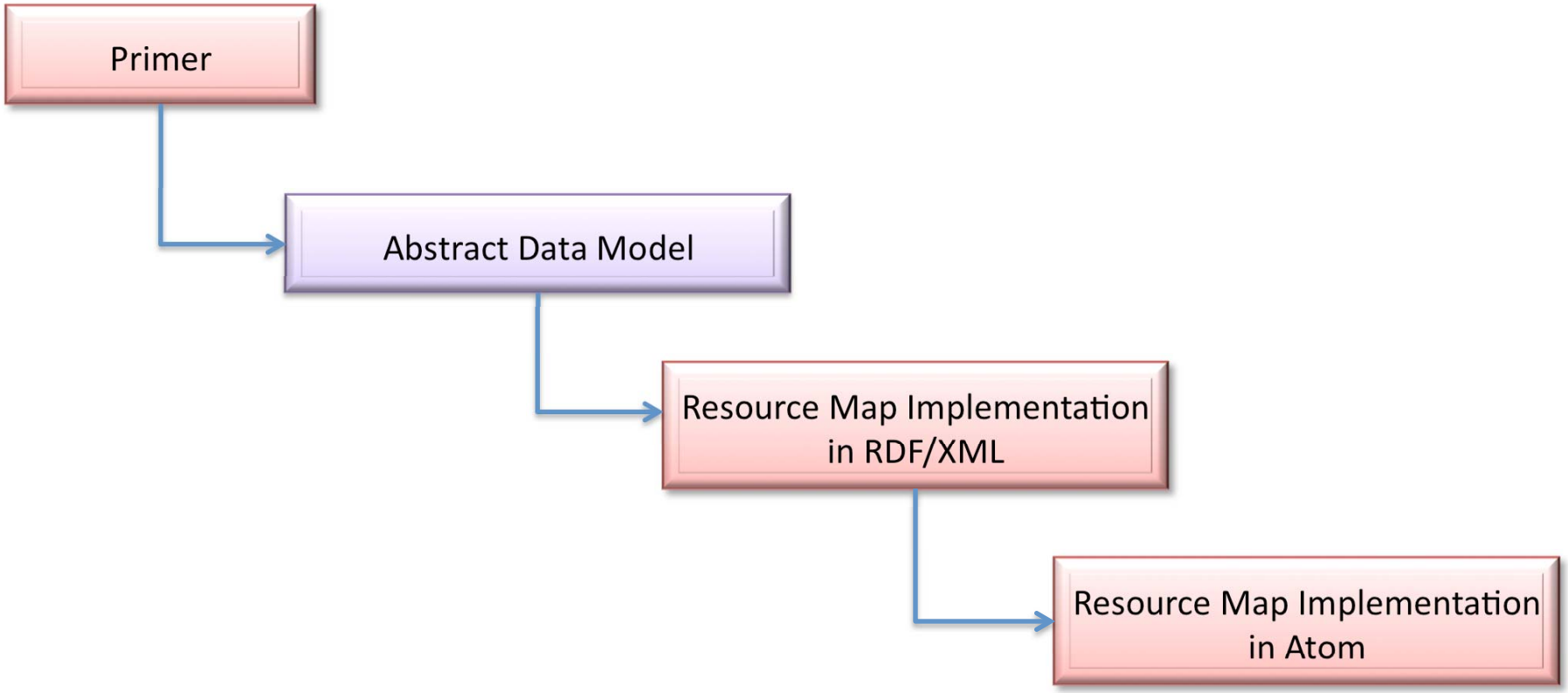



# This Presentation

- OAI-ORE for the most of us
  - Abstract Data Model Basics
    - Aggregations and Aggregated Resources
    - Resource Maps
    - Metadata about Aggregations and Resource Maps
    - Other Resource Map relationships
  - Serializing Resource Maps in RDF/XML
  - Serializing Resource Maps in Atom




# Document Chain



# 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

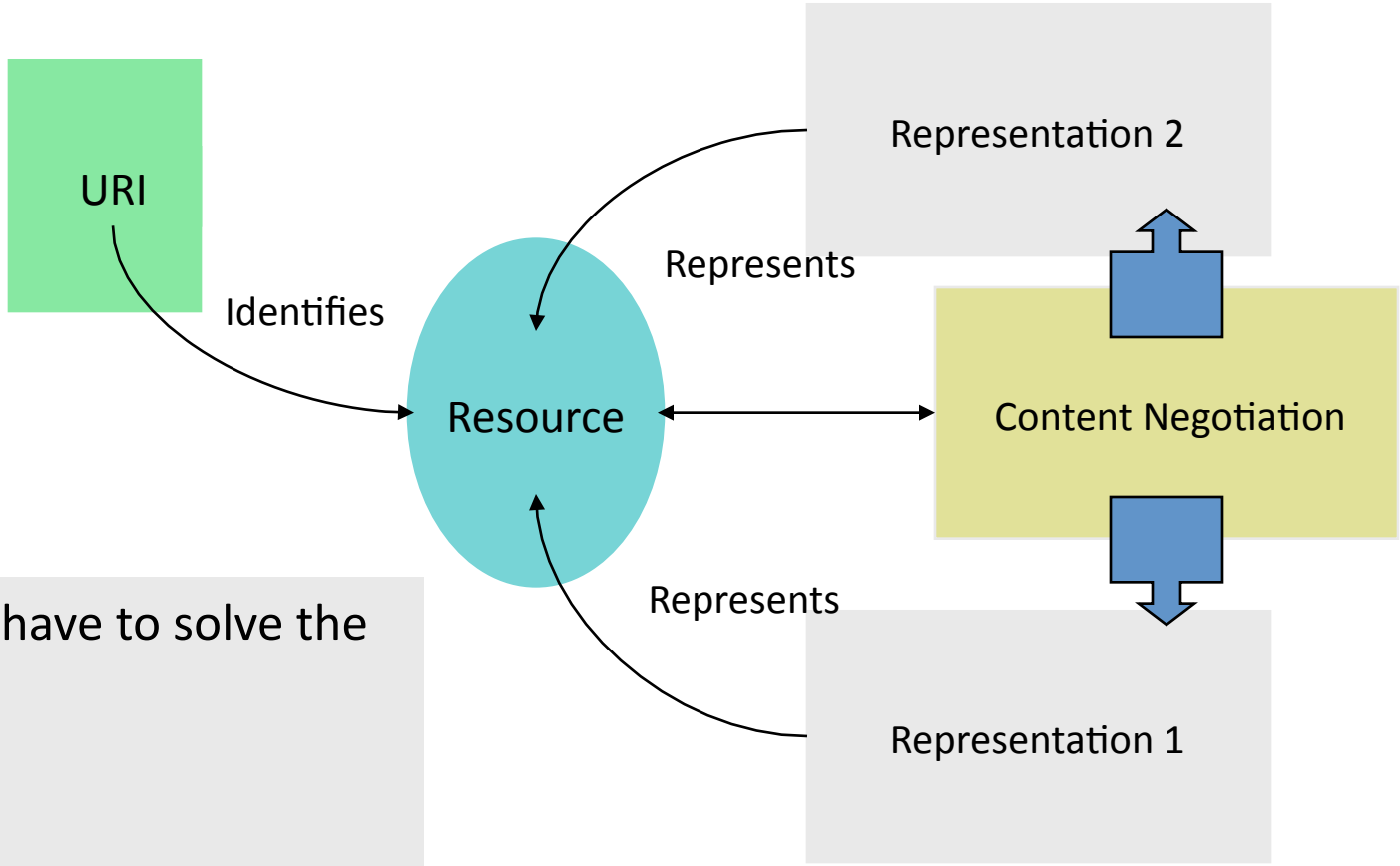



# Requirements of the Model

- 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



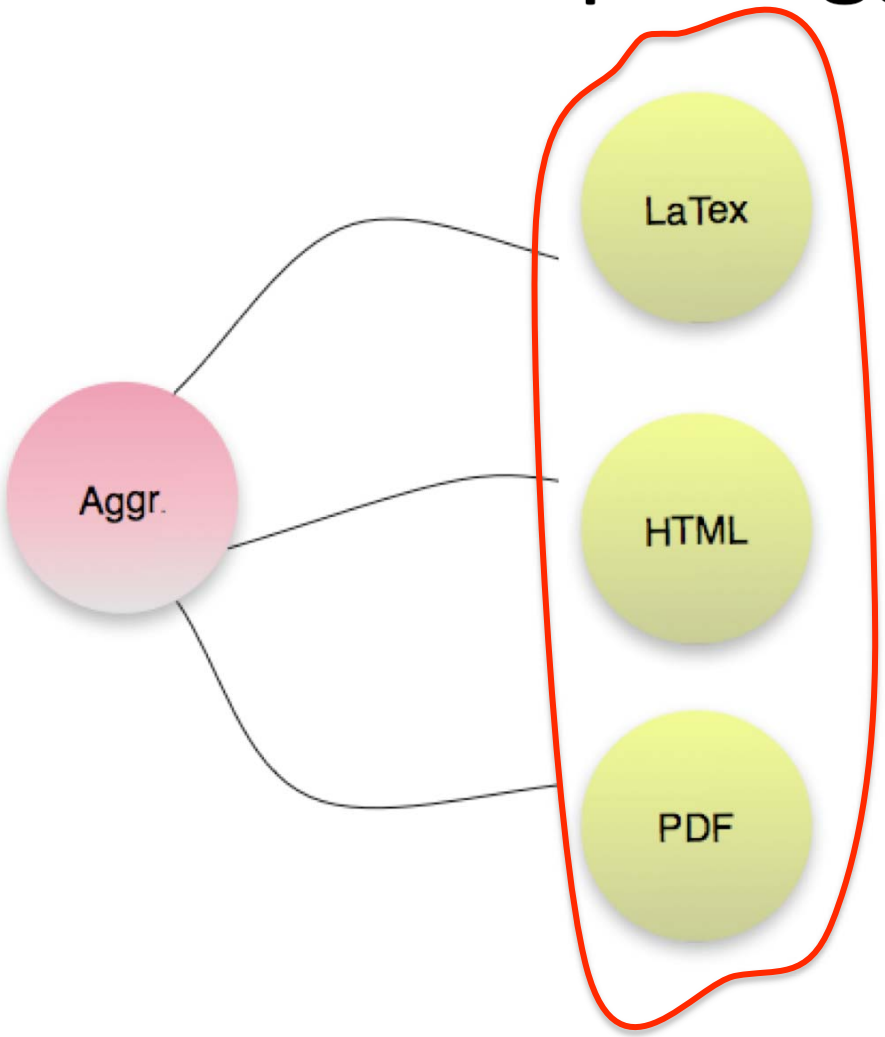

# W3C Web Architecture




So, the tools we have to solve the problem are:

- Resource
- URI
- Representation

# Simple Aggregation

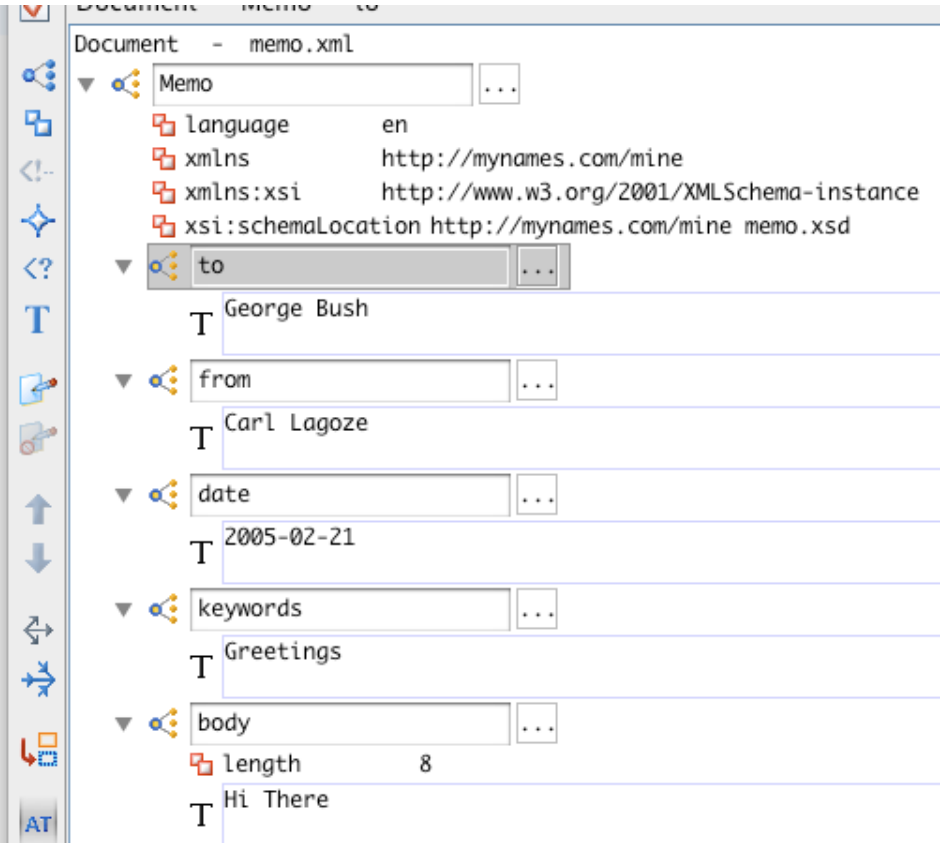



Describe an HTML page, PDF, and Latex as an aggregation

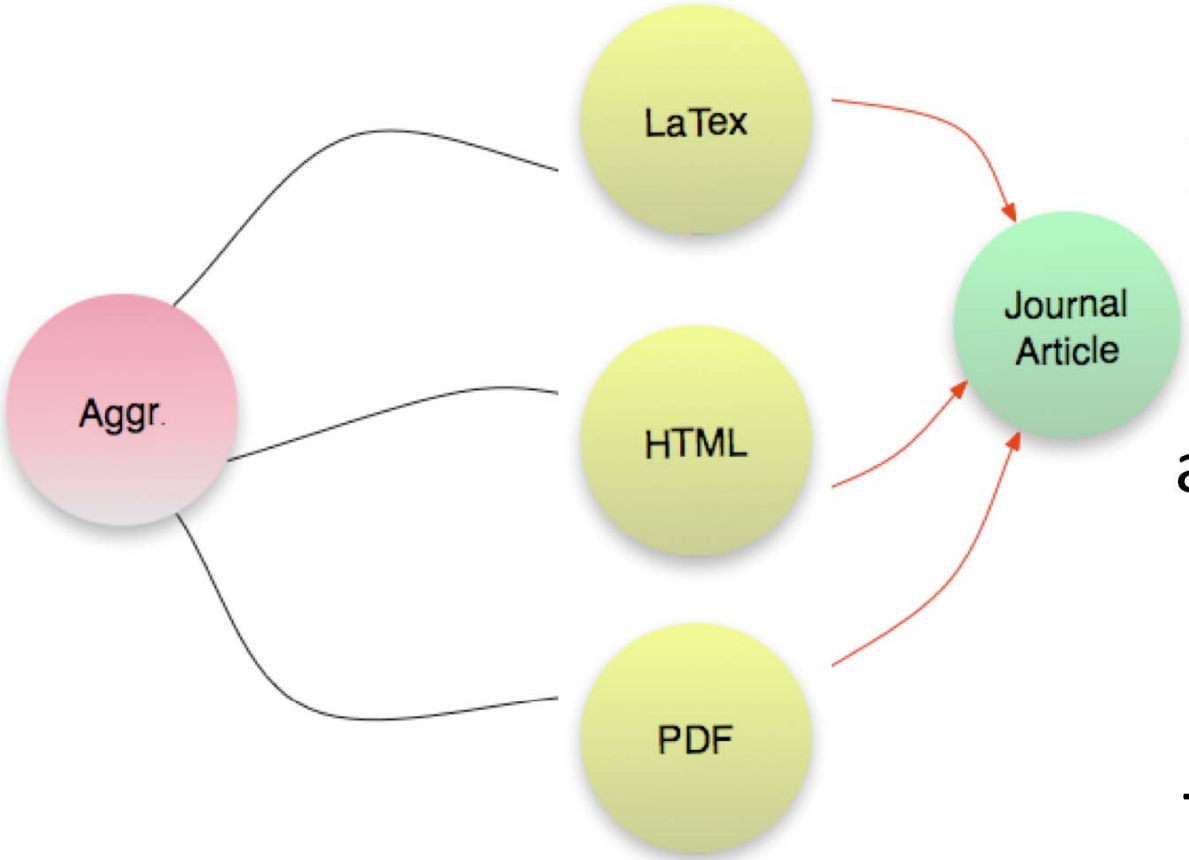
Hierarchical or *Tree* Structure

# Hierarchical Models and XML

```
1  <?xml version="1.0" encoding="UTF-8"?>
2  <Memo xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
3   xsi:schemaLocation="http://mynames.com/mine memo.xsd"
4   xmlns="http://mynames.com/mine"
5   language="en">
6   <to>George Bush</to>
7   <from>Carl Lagoze</from>
8   <date>2005-02-21</date>
9   <keywords>Greetings</keywords>
10  <body length="8">Hi There</body>
11 </Memo>
12
```



# More Complexity

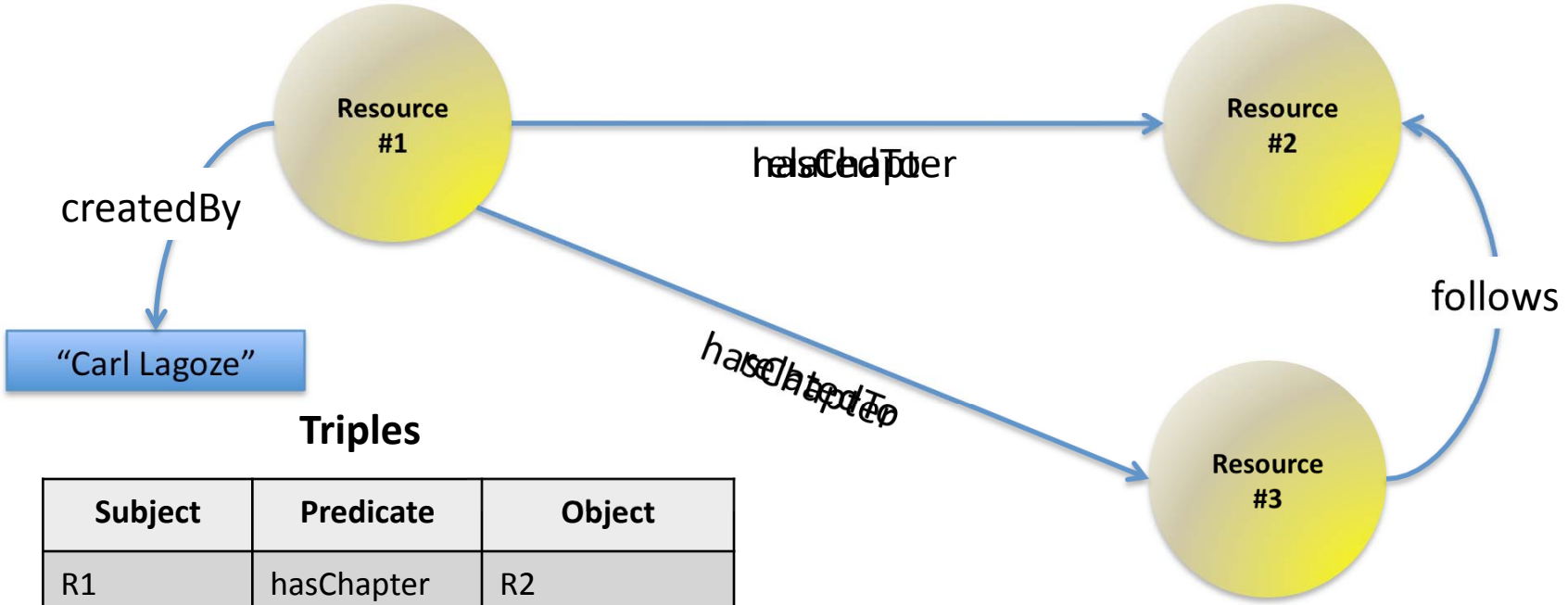



Describe an HTML page, PDF, and LaTeX as an aggregation.

Assert that the PDF and LaTeX are journal articles

*Graph* Structure, Typed Relationships

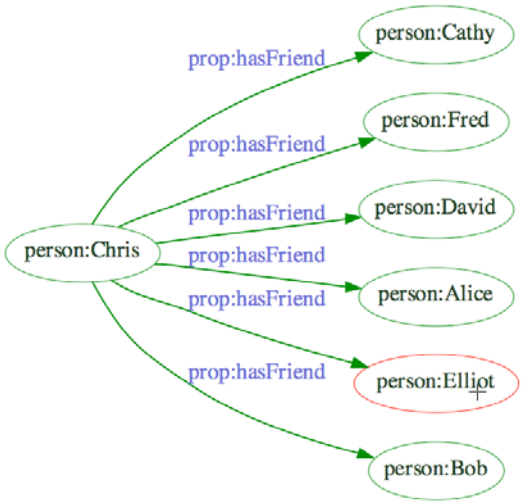
# Resource Description Framework (RDF)




**Triples**

| Subject | Predicate | Object |
|---|---|---|
| R1 | hasChapter | R2 |
| R1 | hasChapter | R3 |
| R3 | follows | R2 |
| R1 | createdBy | "Carl Lagoze" |

# Multiple serializations




```
<rdf:RDF
    xmlns:j.0="prop:h"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
  <rdf:Description rdf:about="person:Chris">
    <j.0:asFriend rdf:resource="person:Cathy"/>
    <j.0:asFriend rdf:resource="person:Fred"/>
    <j.0:asFriend rdf:resource="person:David"/>
    <j.0:asFriend rdf:resource="person:Alice"/>
    <j.0:asFriend rdf:resource="person:Elliot"/>
    <j.0:asFriend rdf:resource="person:Bob"/>
  </rdf:Description>
</rdf:RDF>
```

```
@prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix :         <#> .

<person:Chris>
      <prop:hasFriend> <person:David> , <person:Elliot> , <person:Bob> , <
```

```
<person:Chris> <prop:hasFriend> <person:Cathy> .
<person:Chris> <prop:hasFriend> <person:Fred> .
<person:Chris> <prop:hasFriend> <person:David> .
<person:Chris> <prop:hasFriend> <person:Alice> .
<person:Chris> <prop:hasFriend> <person:Elliot> .
<person:Chris> <prop:hasFriend> <person:Bob> .
```

Open Archives Initiative
Object Reuse and Exchange

# ORE Specification - Abstract Data Model

2 June 2008

***Note: This document is beta and subject to change before final release. 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 beta status. Please comment via the OAI-ORE Google Group.***

**This version:**
http://www.openarchives.org/ore/0.9/datamodel
**Latest version:**
http://www.openarchives.org/ore/datamodel
**Previous version:**
http://www.openarchives.org/ore/0.3/datamodel

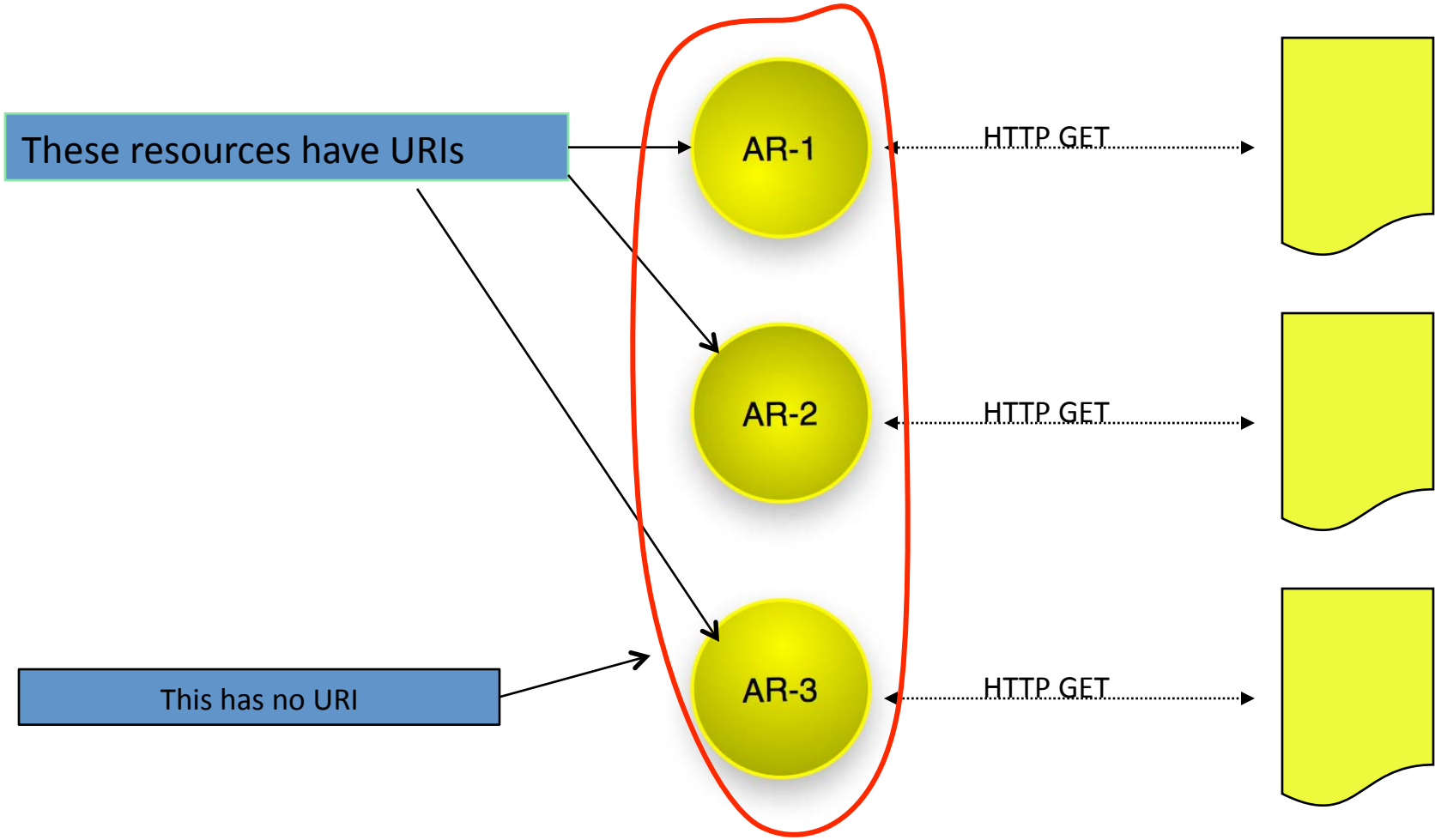
**Editors (OAI Executive)**
Carl Lagoze, Cornell University Information Science
Herbert Van de Sompel, Los Alamos National Laboratory
**Editors (ORE Technical Committee)**
Pete Johnston, Eduserv Foundation
Michael Nelson, Old Dominion University
Robert Sanderson, University of Liverpool
Simeon Warner, Cornell University Information Science

## Abstract

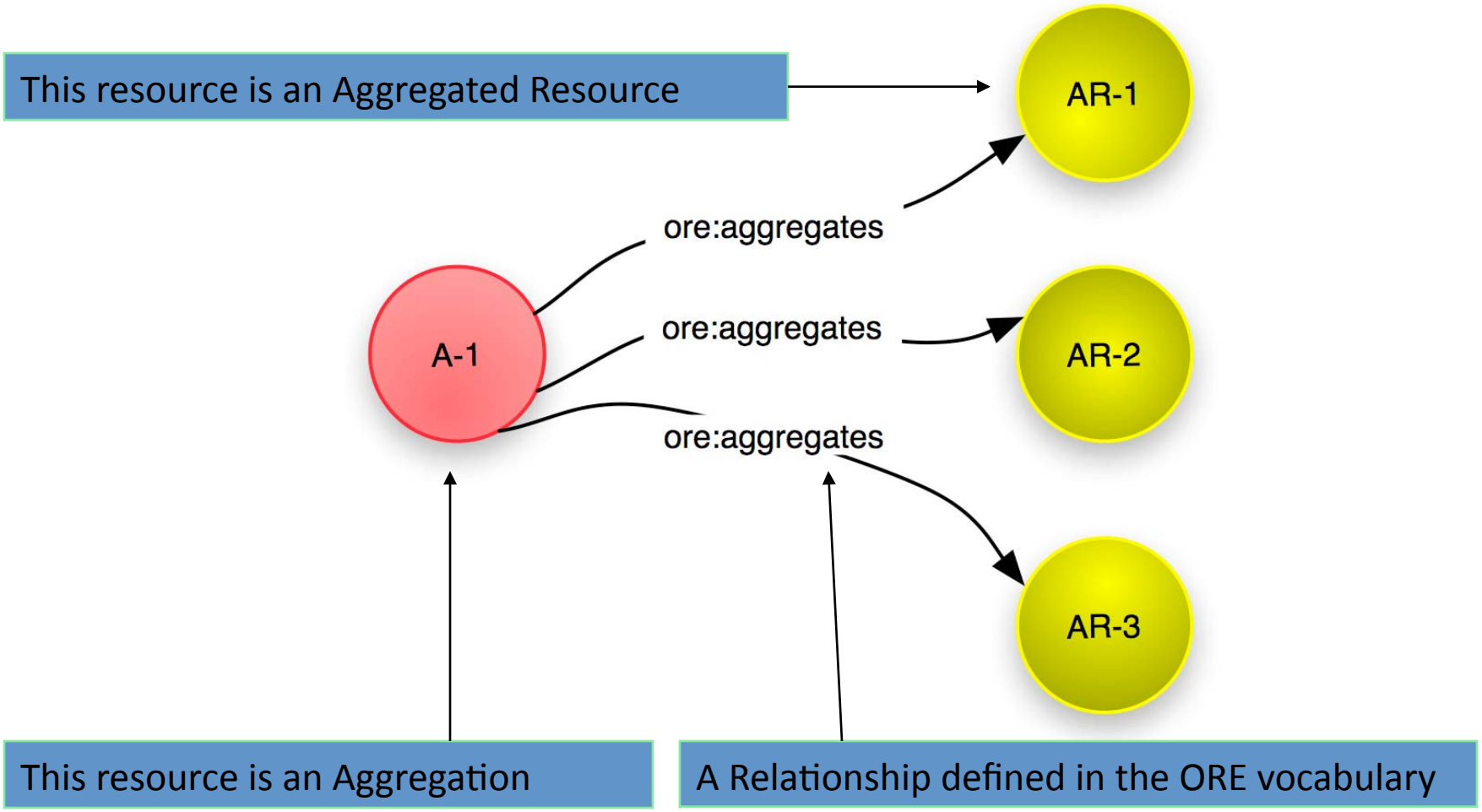
Open Archives Initiative Object Reuse and Exchange (OAI-ORE) defines standards for the description and exchange of aggregations of Web resources. This document describes the abstract data model that is the foundation for these standards. This model is conformant with the Architecture of the World Wide Web [Web Architecture] and leverages Named Graphs [Named Graph] as a mechanism for encapsulating RDF descriptions [RDF Concepts] about aggregations. This specification is one of several documents comprising the OAI-ORE specification and user guide.



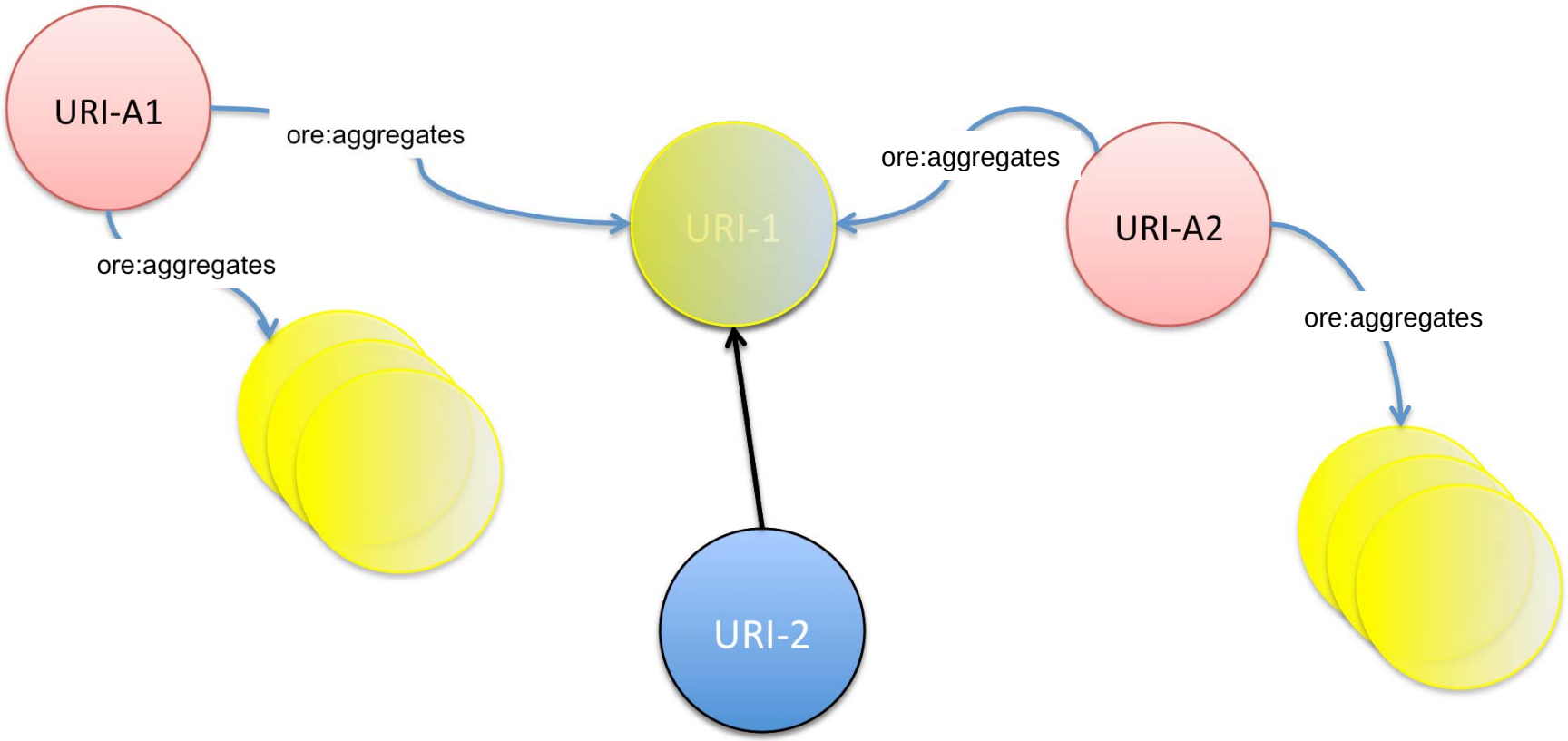

# The starting point: bringing some resources together



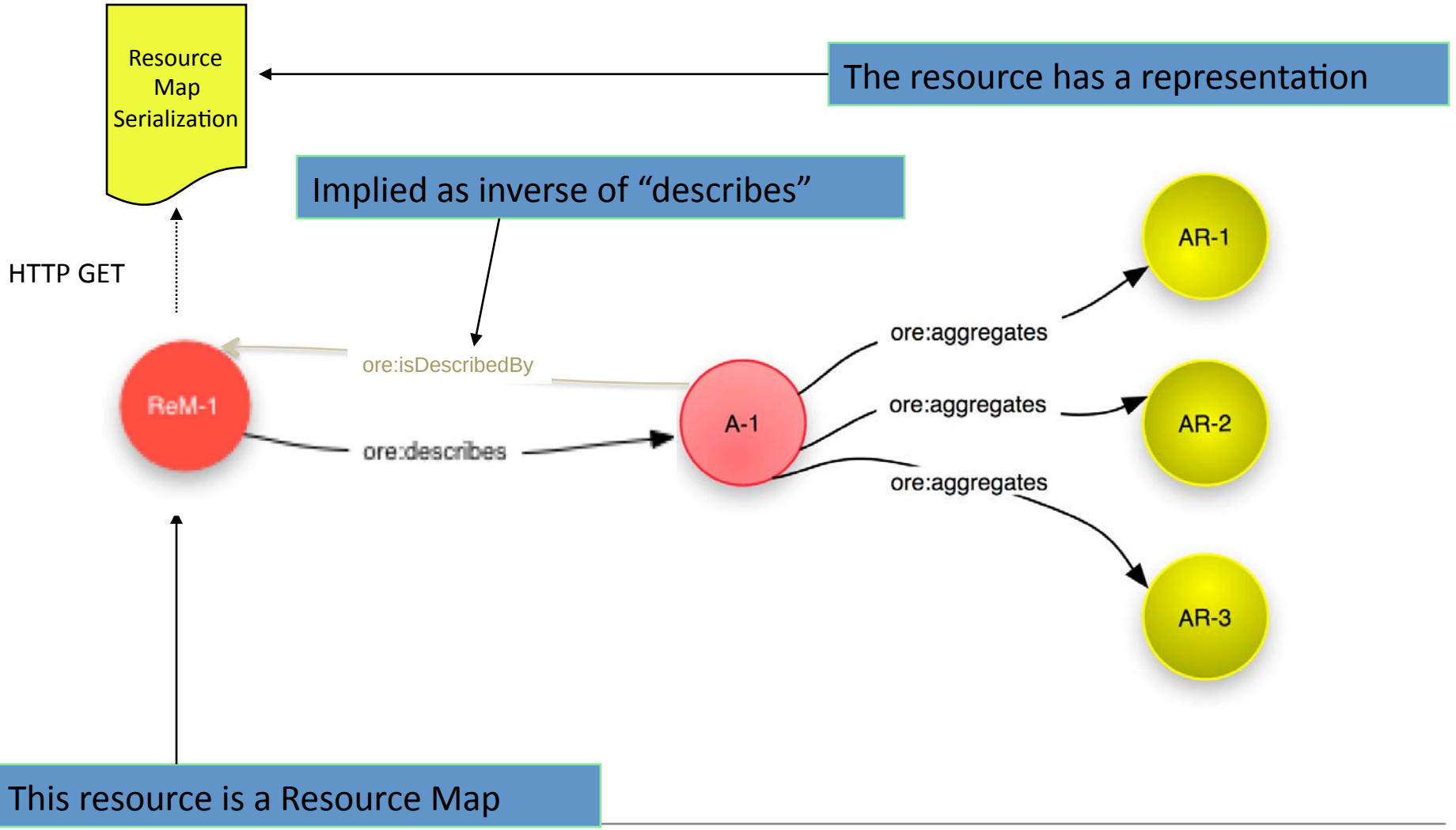
# Aggregation: Resource that is a set of resources



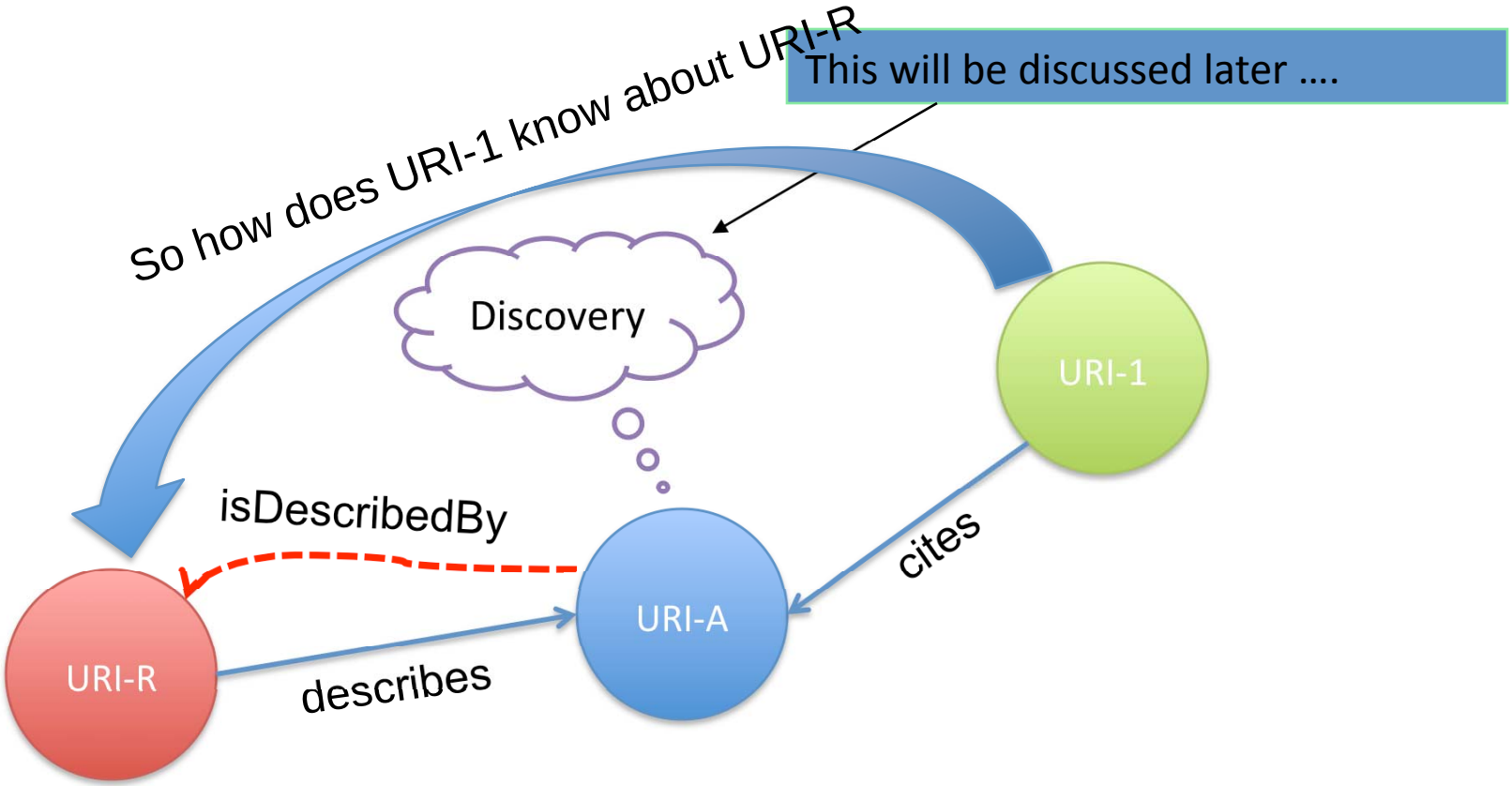
# An Aggregated Resource is just a Resource



# Describe an Aggregation: Resource Map



# Relationship between Aggregation and Resource Map

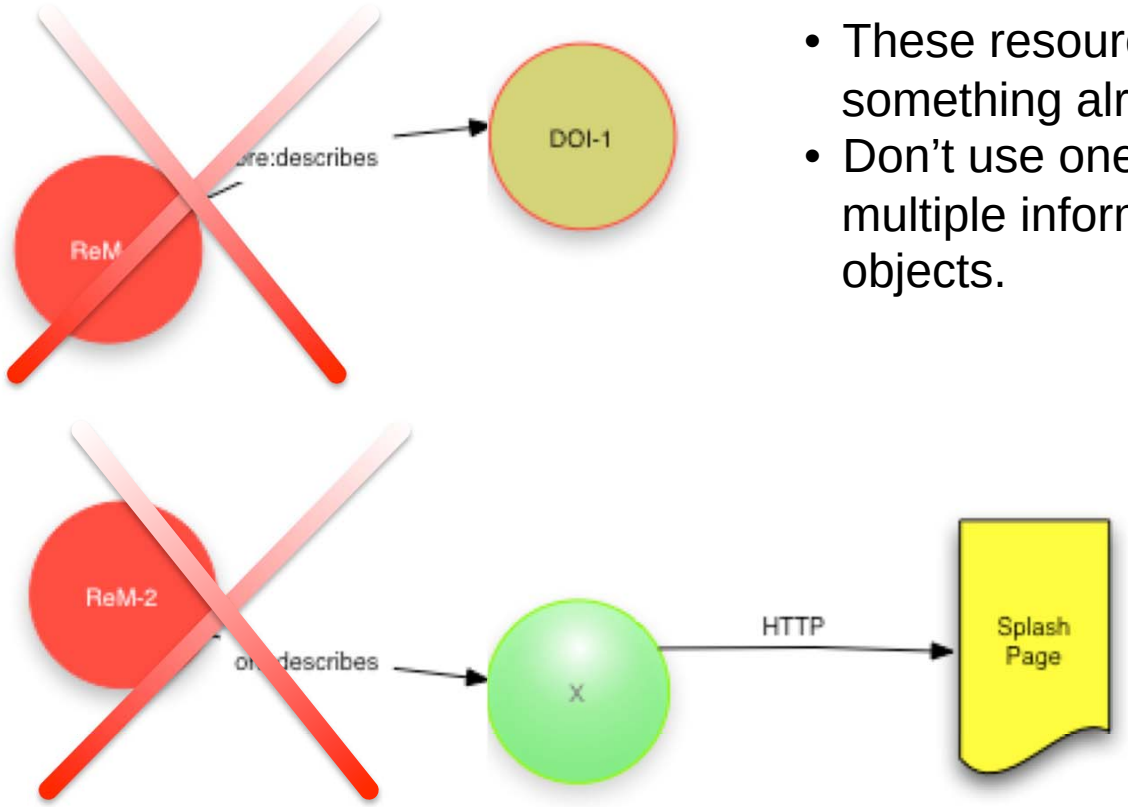


# Relationship between Aggregation and Resource Map

- An Aggregation is a Resource with a URI
- A Resource Map is a Resource with a URI
- A Resource Map asserts (identifies) and describes one Aggregation
  - A Resource is an Aggregation due to an assertion by (at least) one Resource Map
  - A Resource Map MUST have one Representation

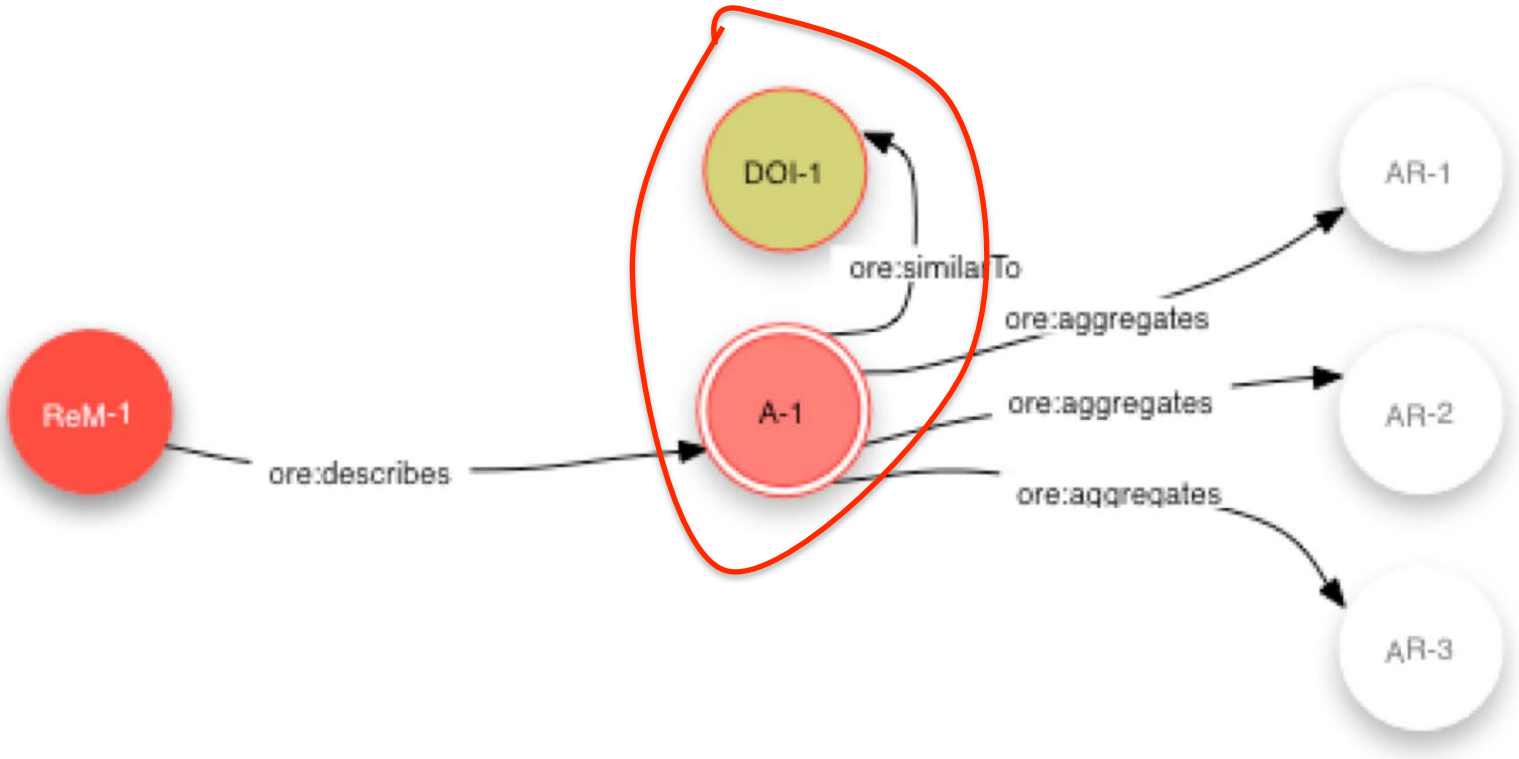



# Don’t overload URI-A




- These resources “mean” something already.
- Don’t use one URI for multiple information objects.

# The right way …

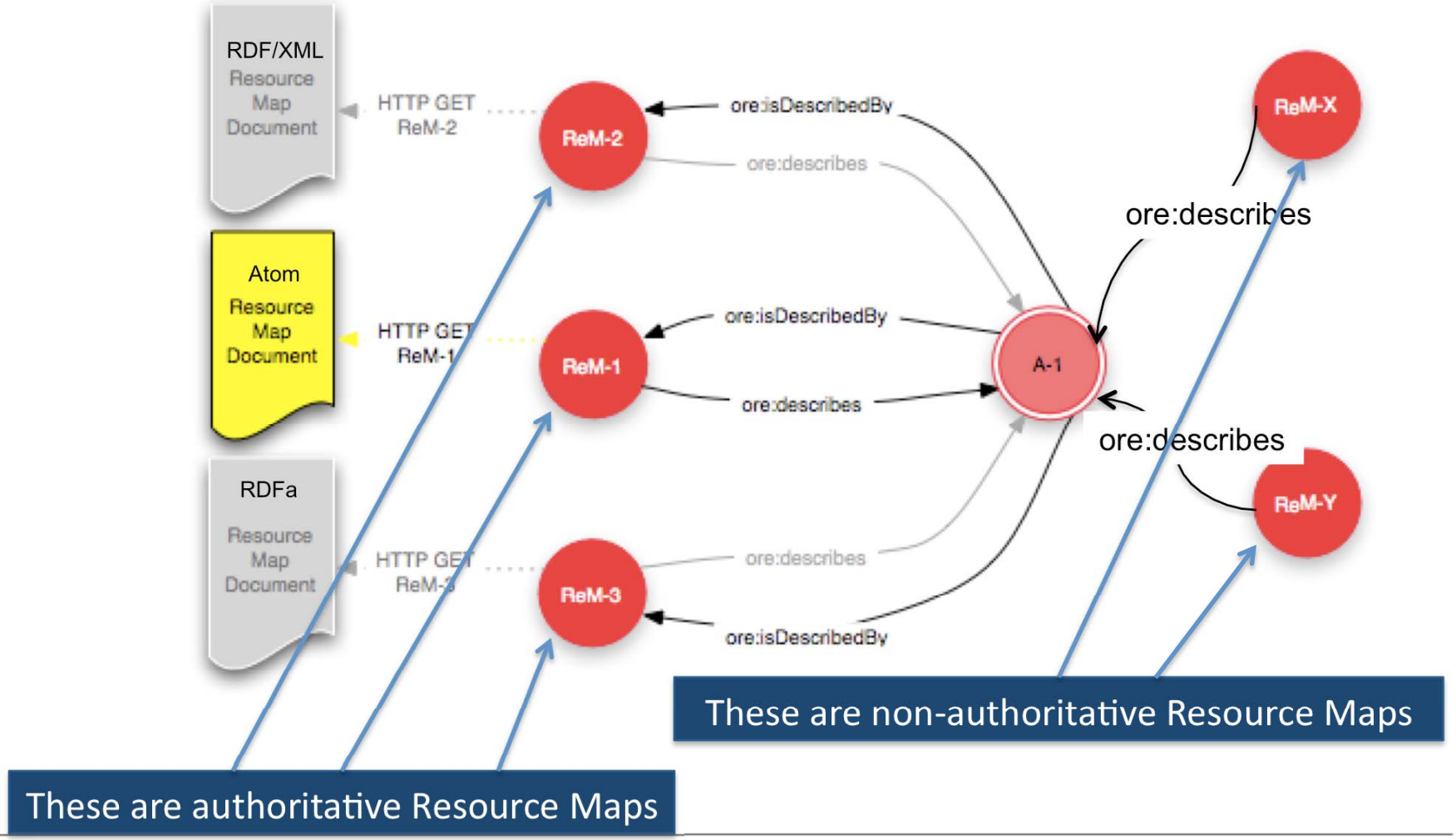


# Multiple Resource Maps

- An Aggregation MAY be asserted and described by multiple Resource Maps
- The purpose of multiple Resource Maps is to provide descriptions of the Aggregation in multiple serializations (e.g., Atom, RDF/XML, RDFa, etc.)
- Each Resource Map MUST have only one representation



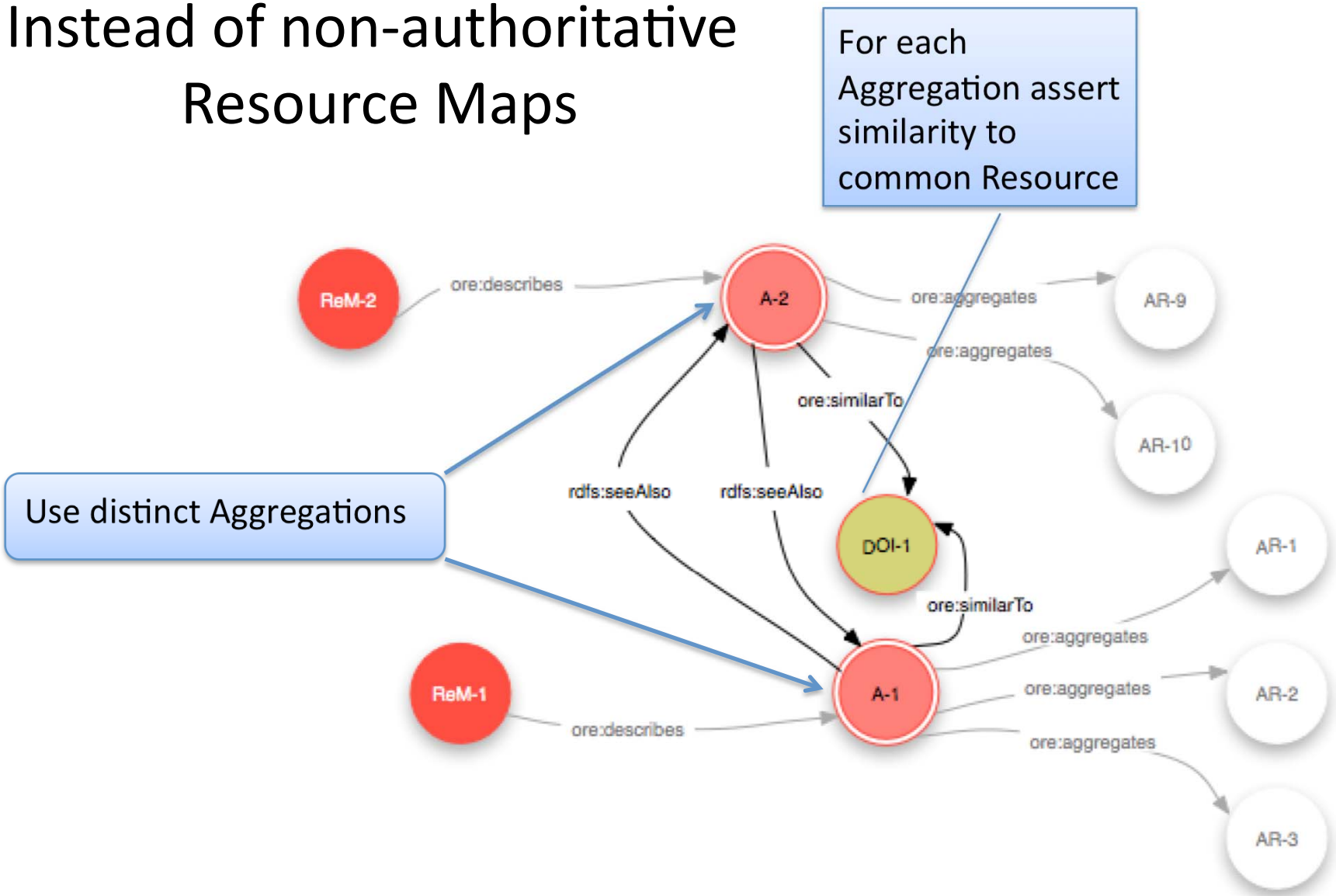

# Multiple Resource Maps



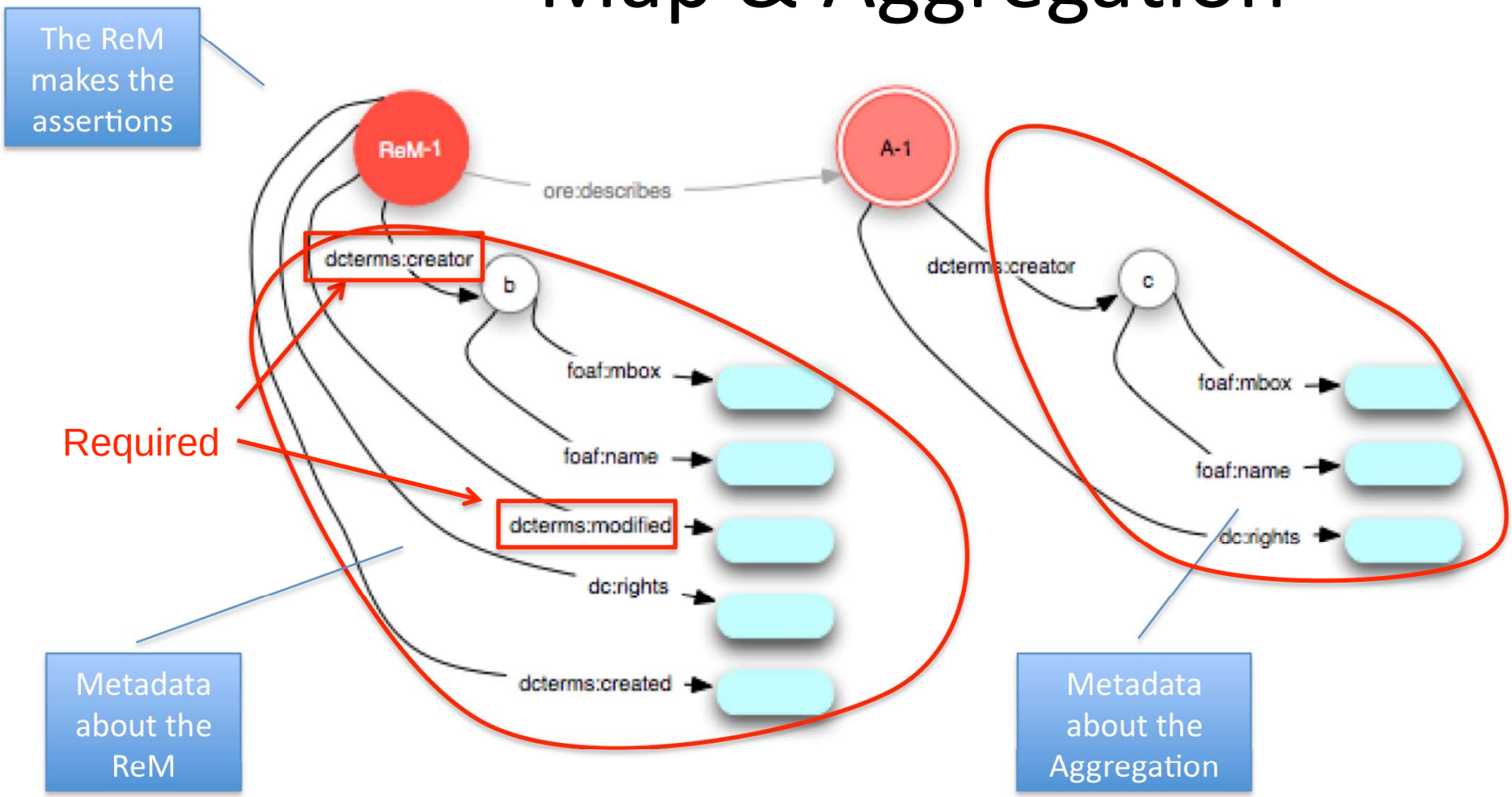
# Authoritative vs. Non-Authoritative Resource Maps

- Authoritative
  - Created by same authority (usually)
  - MUST be minimally equivalent (same Aggregated Resources and Proxies)
  - SHOULD assert mutual existence
- Non-authoritative
  - Best practice is to not create them
  - Assert your own Aggregation instead
  - Use rdfs:seeAlso to assert relationship between two Aggregations

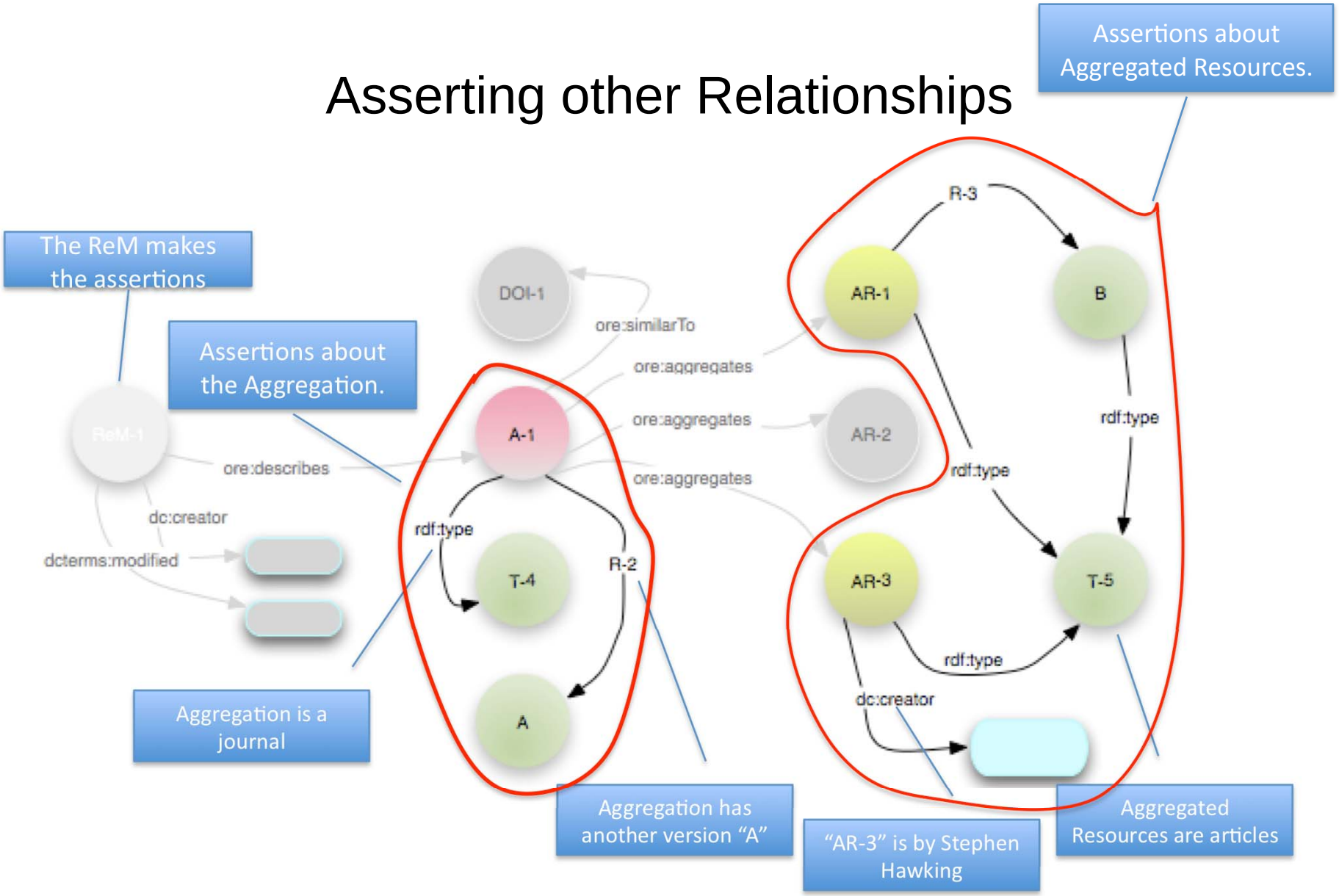
# Instead of non-authoritative Resource Maps



# Metadata about the Resource Map & Aggregation



# Asserting other Relationships



# 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 visiting Pittsburgh
- 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




# Open Archives Initiative Object Reuse & Exchange

## Basics:

## Serializing the Model




# 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




Open Archives Initiative
Object Reuse and Exchange

# ORE User Guide - Resource Map Implementation in RDF/XML

2 June 2008

***Note: This document is beta and subject to change before final release. 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 beta status. Please comment via the OAI-ORE Google Group.***

**This version:**
http://www.openarchives.org/ore/0.9/rdfxml
**Latest version:**
http://www.openarchives.org/ore/rdfxml
**Previous version:**
http://www.openarchives.org/ore/0.3/rdfsyntax

**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. OAI-ORE introduces the notion of a Resource Map, a named RDF Graph [RDF Concepts] which describes the Aggregation, the Aggregated Resources of which it is composed, and the relationships between them (and/or the relationships between these and other resources). Since a Resource Map is an RDF Graph, it can be serialized using any RDF syntax. This document outlines the use of one such syntax for the serialization of Resource Maps: RDF/XML [RDF/XML]. A companion document outlines the serialization of Resource Maps in RDFa [RDFa Syntax]. This is one of several documents comprising the OAI-ORE specification and user guide.



```
<?xml version="1.0" encoding="UTF-8" ?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:ore="http://www.openarchives.org/ore/terms/" xmlns:dcterms="http://purl.org/dc/terms/"
    xmlns:foaf="http://xmlns.com/foaf/0.1/">
    <rdf:Description rdf:about="http://arxiv.org/astro-ph/0601007/foo.rdf">
        <ore:describes rdf:resource="http://arxiv.org/astro-ph/0601007#aggregation"/>
        <dcterms:creator rdf:resource="http://my.example.org/agents/AgencyX"/>
        <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#date"
        >2008-02-12</dcterms:modified>
    </rdf:Description>
    <rdf:Description rdf:about="http://arxiv.org/astro-ph/0601007#aggregation">
        <ore:isDescribedBy rdf:resource="http://arxiv.org/astro-ph/0601007/foo.atom"/>
        <dcterms:creator rdf:parseType="Resource">
            <foaf:name>Hui Li</foaf:name>
        </dcterms:creator>
        <ore:aggregates rdf:resource="http://arxiv.org/ps/astro-ph/0601007"/>
        <ore:aggregates rdf:resource="http://arxiv.org/pdf/astro-ph/0601007"/>
        <ore:aggregates rdf:resource="http://arxiv.org/e-print/astro-ph/0601007"/>
    </rdf:Description>
    <rdf:Description rdf:about="http://my.example.org/agents/AgencyX">
        <foaf:name>The Best ReM Agent</foaf:name>
    </rdf:Description>
</rdf:RDF>
```

Resource Map

Aggregation

Aggregated Resources




Open Archives Initiative
Object Reuse and Exchange

# ORE User Guide - Resource Map Implementation in Atom

2 June 2008

***Note: This document is beta and subject to change before final release. 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 beta status. Please comment via the OAI-ORE Google Group.***

**This version:**
http://www.openarchives.org/ore/0.9/atom-implementation
**Latest version:**
http://www.openarchives.org/ore/atom-implementation
**Previous version:**
http://www.openarchives.org/ore/0.3/atom-implementation

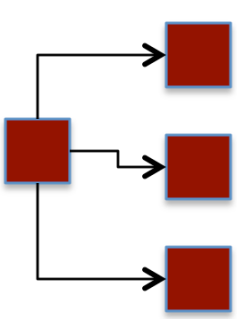
**Editors (OAI Executive)**
Carl Lagoze, Cornell University Information Science
Herbert Van de Sompel, Los Alamos National Laboratory
**Editors (ORE Technical Committee)**
Pete Johnston, Eduserv Foundation
Michael Nelson, Old Dominion University
Robert Sanderson, University of Liverpool
Simeon Warner, 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. OAI-ORE introduces the notion of Resource Maps that describe an Aggregation. 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 relationships pertaining to the Aggregation and its Aggregated Resources. Each Aggregation may be described by one or more Resource Maps, each of which must have exactly one representation that is a serialization of the Resource Map according to a specific format. A detailed examination of the



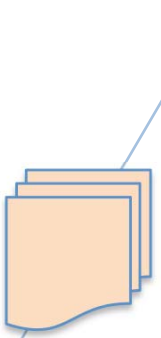

# Syndication

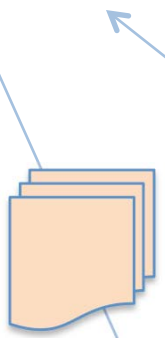


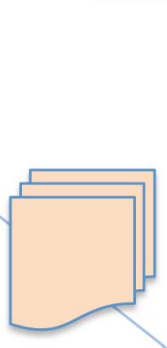
Simple
Compound
Document

RSS
- 0.x
- 1.0
- 2.0

Atom







# Atom

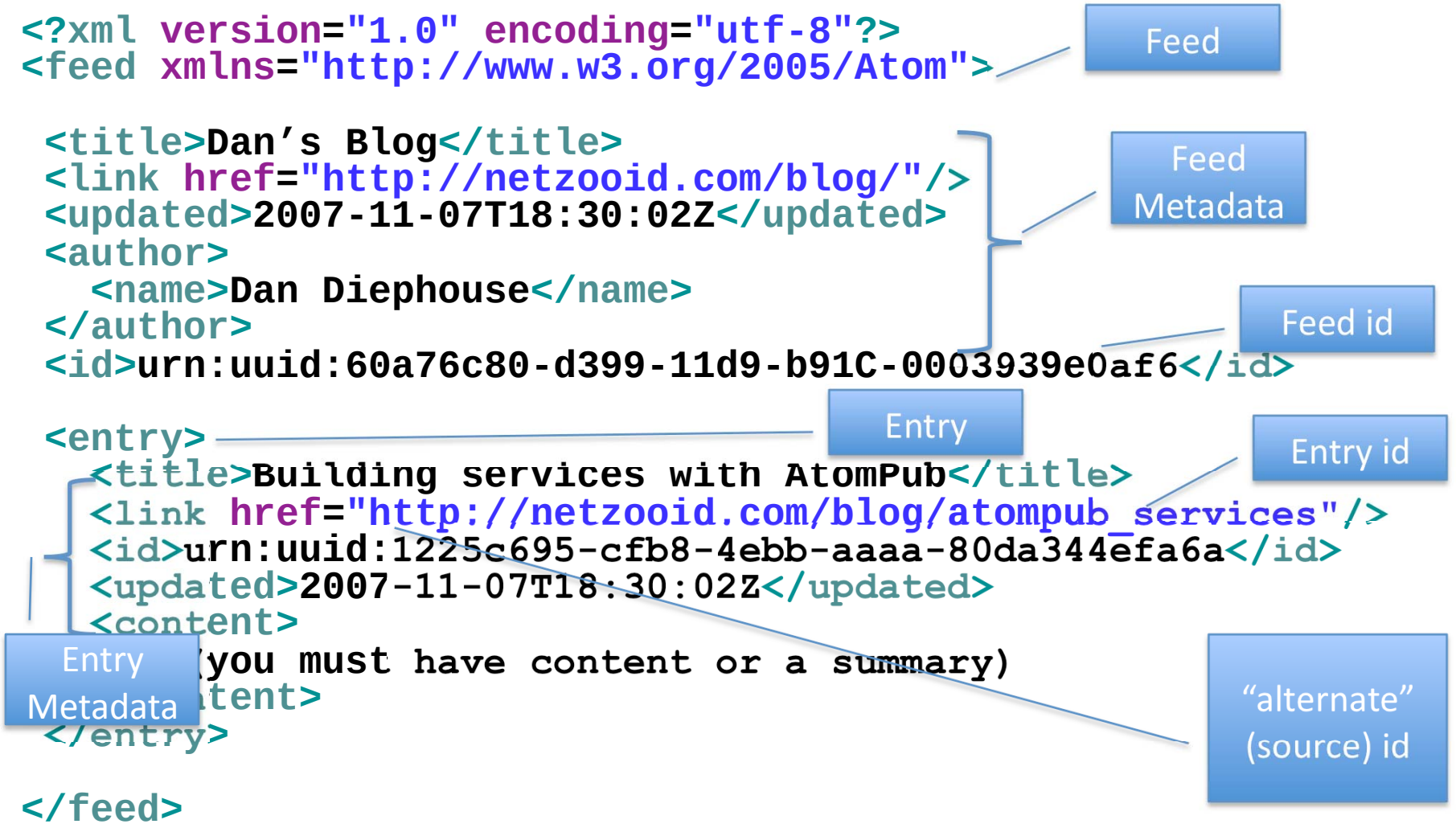
- Attempt to rationalize RSS 1.x, 2.x divergence
- Encoding is up-to-date with current XML standards
  - namespaces
  - Schema
- Robust content model
  - Distinguishes between metadata and content (plain text, HTML, base-64 binary)
- Well-defined extensibility model
- IETF FRC 4287
  - http://www.ietf.org/rfc/rfc4287




# Structure of Feed

- **Feed**
  - ID
  - Author
  - Link
  - Title
  - Updated
  - *
  - **Entry**
    - ID
    - Updated
    - Link
    - Summary
    - Content
    - *

# Simple Atom Feed




```
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns="http://www.w3.org/2005/Atom">

  <title>Dan's Blog</title>
  <link href="http://netzooid.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://netzooid.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)
    tent>
  </entry>

</feed>
```

# 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 only when necessary

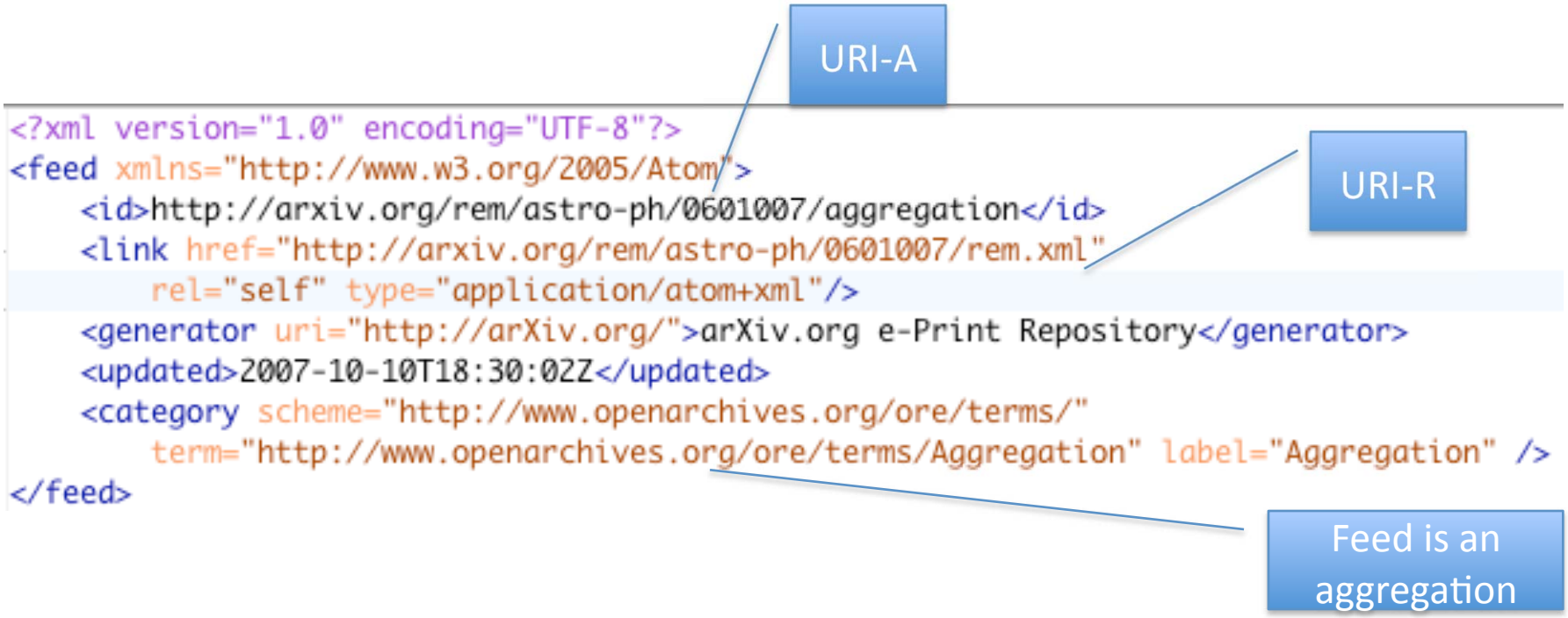



# Mapping ORE Model to Atom

| anoORE | Atom |
|---|---|
| *Aggregation* | *Feed* |
| URI-A | Feed <id> |
| URI-R | <link href="URI" rel="self" |
| ore:similarTo | <link href="URI" rel="related" |
| Aggregation Properties/Metadata | Feed metadata |
| *Aggregated Resource* | *Entry* |
| URI-AR | <link href="URI" rel="alternate" |
| Aggregated Resource Properties/Metadata | Entry metadata |




# Atom/ORE Skeleton




```
<?xml version="1.0" encoding="UTF-8"?>
<feed xmlns="http://www.w3.org/2005/Atom">
    <id>http://arxiv.org/rem/astro-ph/0601007/aggregation</id>
    <link href="http://arxiv.org/rem/astro-ph/0601007/rem.xml"
        rel="self" type="application/atom+xml"/>
    <generator uri="http://arXiv.org/">arXiv.org e-Print Repository</generator>
    <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>
```

# Adding Entries

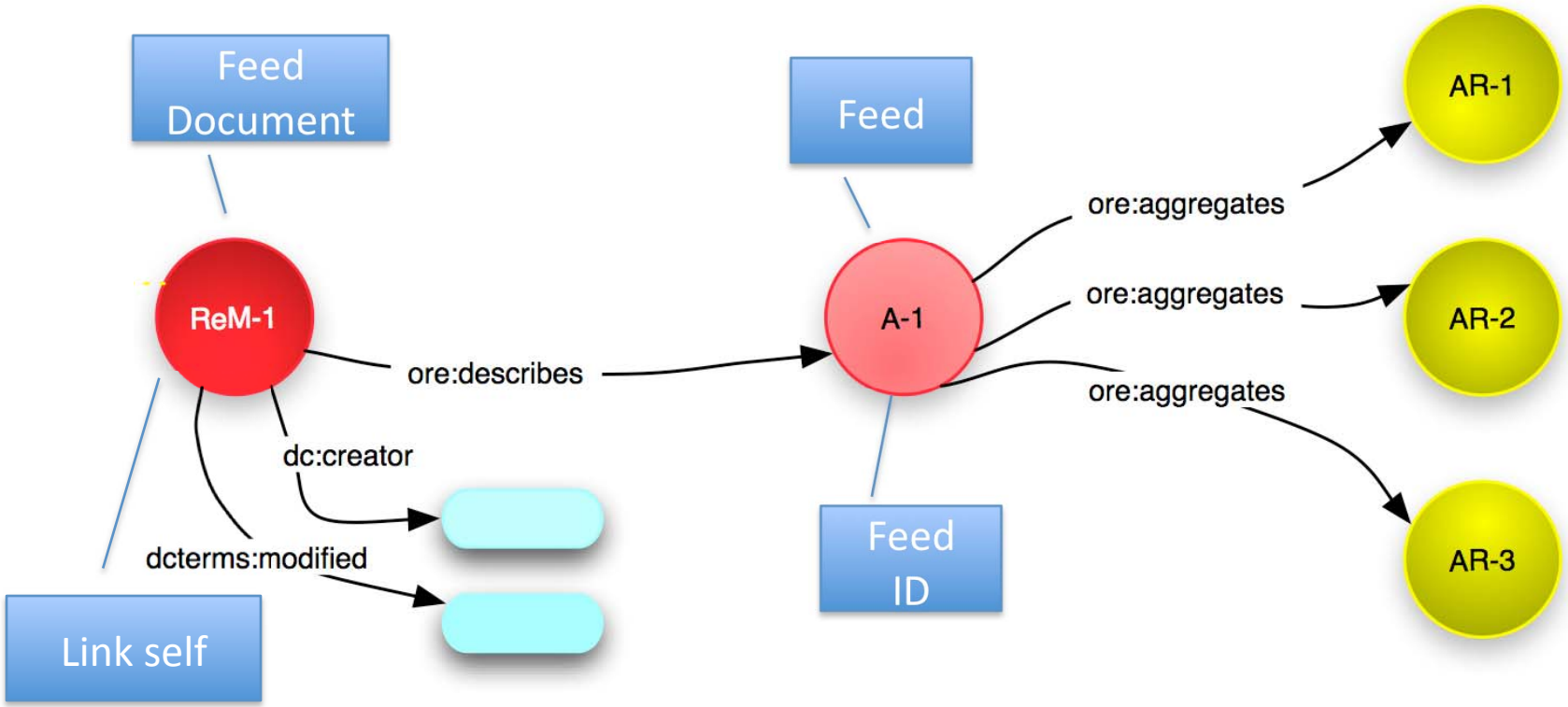
URI-AR

```
<?xml version="1.0" encoding="UTF-8"?>
<feed xmlns="http://www.w3.org/2005/Atom">
    <id>http://arxiv.org/rem/astro-ph/0601007/aggregation</id>
    <link href="http://arxiv.org/rem/astro-ph/0601007/rem.xml" rel="self" [1 line]
    <generator uri="http://arXiv.org/">arXiv.org e-Print Repository</generator>
    <updated>2007-10-10T18:30:02Z</updated>
    <category scheme="http://www.openarchives.org/ore/terms/" [1 line]
    <entry>
        <id>tag:arxiv.org,2007:astro-ph/0601007v2:ps</id>
        <link href="http://arxiv.org/ps/astro-ph/0601007" rel="alternate"
            type="application/postscript"/>
    </entry>
    <entry>
        <id>tag:arxiv.org,2007:astro-ph/0601007v2:pdf</id>
        <link href="http://arxiv.org/pdf/astro-ph/0601007" rel="alternate"
            type="application/pdf"/>
    </entry>
    <entry>
        <id>tag:arxiv.org,2007:astro-ph/0601007v2:e-print</id>
        <link href="http://arxiv.org/e-print/astro-ph/0601007" rel="alternate"/>
    </entry>
</feed>
```





# Data Model <-> Atom



# Looking to the future with Atom

- Atom Publishing Protocol
- SWORD
- Microsoft/Google APIs



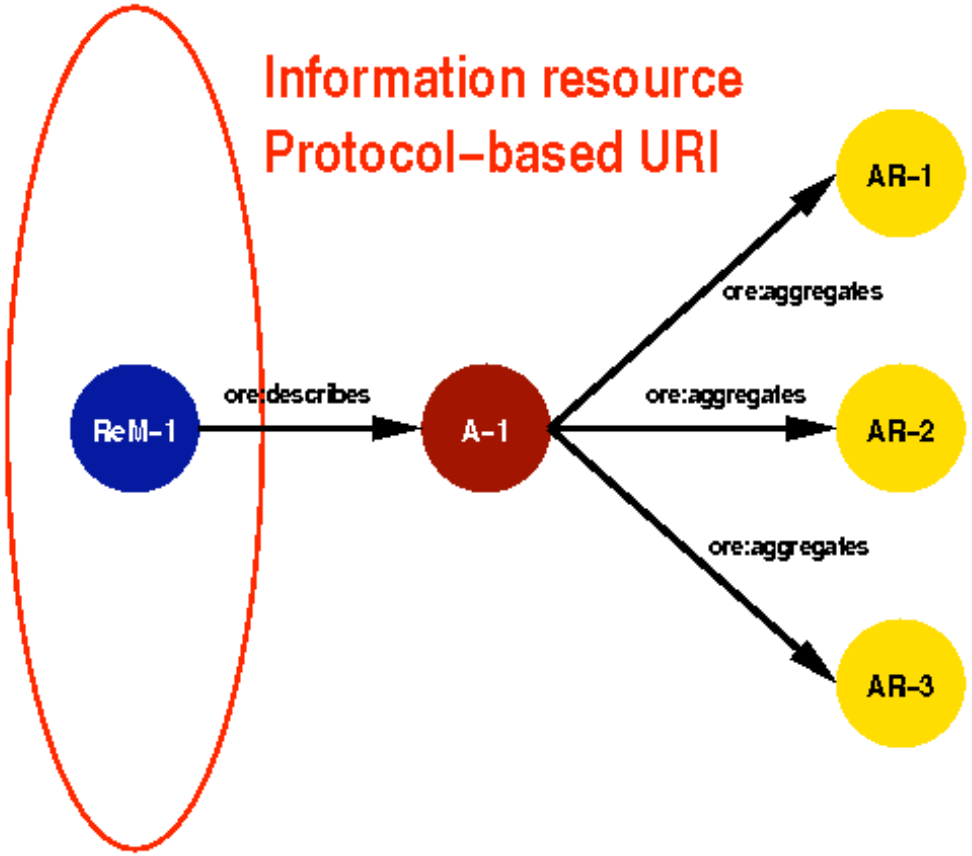

# Open Archives Initiative Object Reuse & Exchange

## HTTP Implementation and Resource Map Discovery

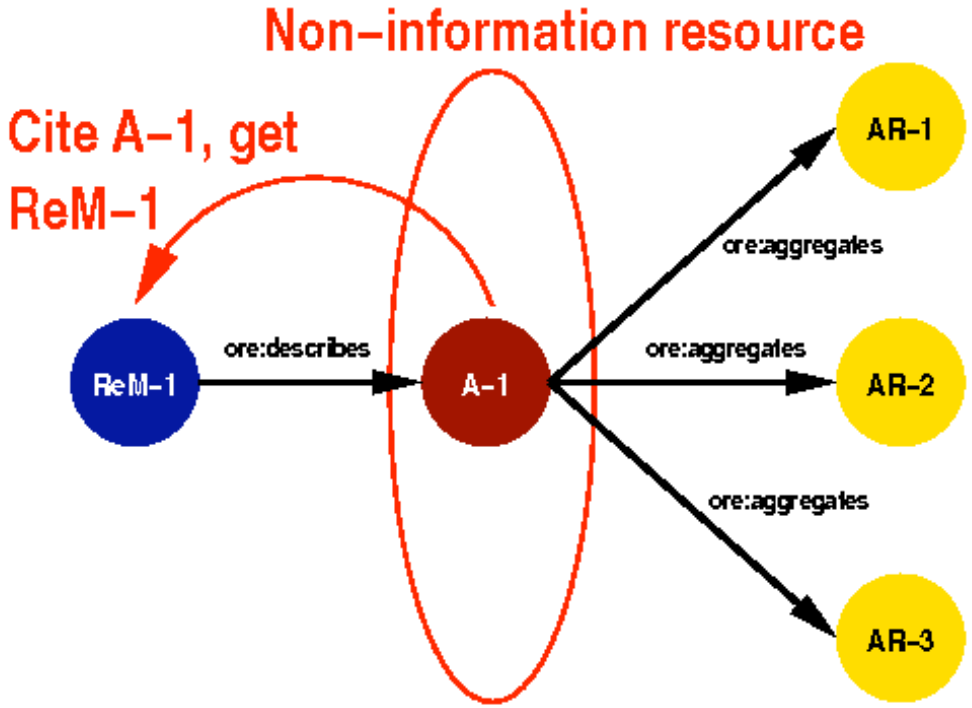



# HTTP Implementation

WHY? -- The Web is built of http URIs



- Require access to Aggregation URI to redirect to or yield a Resource Map
- Several strategies...



# Cool URIs

- Want simple, stable, manageable URIs
  - stability important for citation
- Certainly no technology baggage (.php, .asp etc.)
- Aggregation URI not tied to format of ReM

```
A-1   = http://example.org/foo
ReM-1 = http://example.org/foo.atom
```




# Two ways to do Cool URIs

## Content Negotiation

- HTTP GET A-1
- Return ReM & URI

Status: 200 OK
Content-Location: ReM-1

<?xml version="1.0" ...>
...ReM-1 in response...

## 303 Redirection

- HTTP GET A-1
- Redirect to ReM-1

(from linked data world)

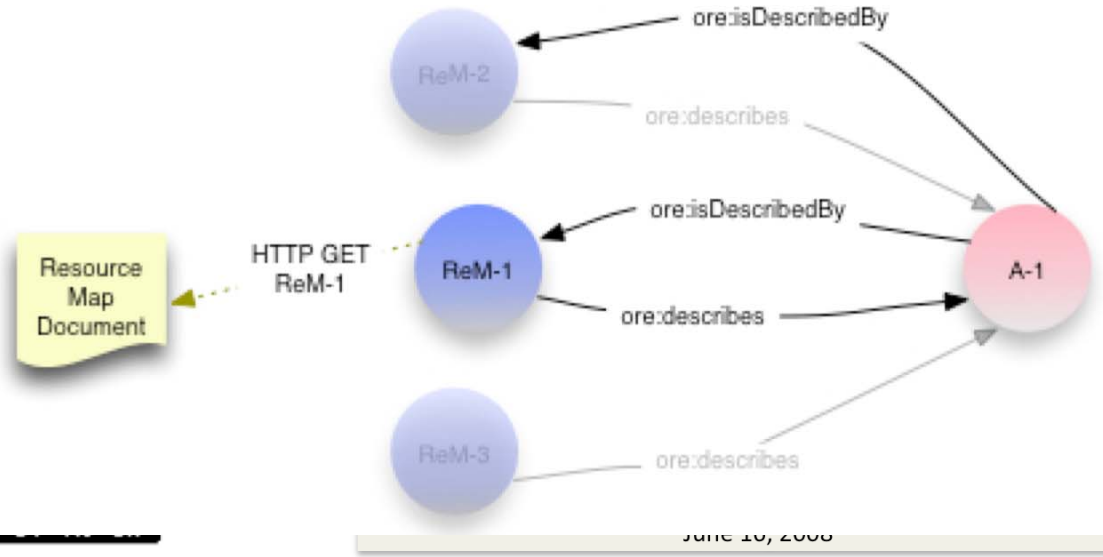
Status: 303 See Other
Location: ReM-1




# Multiple Resource Maps

```
A-1   = http://example.org/foo
ReM-1 = http://example.org/foo.atom
ReM-2 = http://example.org/foo.rdf
```

- Choose one Resource Map as default
  (access from A-1 → authoritative)
- Indicate other ReMs with ore:isDescribedBy
  (chain of authority)



# No Server Support

```
A-1   = http://example.org/foo.atom#aggregation
ReM-1 = http://example.org/foo.atom
```

- Use fragment identifier
- Recommended in CoolURI and Linked Data Specs
- Extension to multiple serializations is ugly




# RDFa and Microformats

- Will discuss just RDFa
- "Splash Page" acts as human description as Resource Map
- "Splash Page" may be Resource Map and an Aggregated Resource (recommend different URIs)
- Resource Map URI MUST be different from Aggregation URI (as usual)




```
<?xml version="1.0" encoding="UTF-8"?>
<html xmlns="http://www.w3.org/1999/xhtml" ... >
<head profile="http://www.w3.org/1999/xhtml/vocab">
  <title>Resource Map http://my.example.org/rem</title>
</head>
<body>
  <div about="http://my.example.org/rem"
       typeof="ore:ResourceMap" class="ResourceMap">
  <h1>Resource Map http://my.example.org/rem</h1>
  <p>Described Aggregation: <a rel="ore:describes"
    href="http://my.example.org/aggregation">
    http://my.example.org/aggregation</a></p>
  <p>Creator: <a rel="dcterms:creator"
    href="http://my.example.org/agtx">Agency X</a>.</p>
  <p>Modification date:
    <span property="dcterms:modified" datatype="xsd:date"
     content="2008-02-12">12 February 2008</span>.</p>
  </div>
</body>
</html>
```




# RFDa – CoolURI & #

- No big change

- CoolURI:

```
A-1    = http://example.org/foo
ReM-1  = http://example.org/foo.html   (+RDFa)
ReM-2  = http://example.org/foo.atom
```

- Fragment identifier:

```
A-1    = http://example.org/foo.html#aggregation
ReM-1  = http://example.org/foo.html   (+RDFa)
```





# Different Splash Page and Resource Map URIs

- Use fragment id to distinguish ReM:

```
S-1   = http://example.org/foo.html      (+RFDa)
A-1   = http://example.org/foo
ReM-1 = http://example.org/foo.html#rem  (+RDFa)
ReM-2 = http://example.org/foo.atom
```

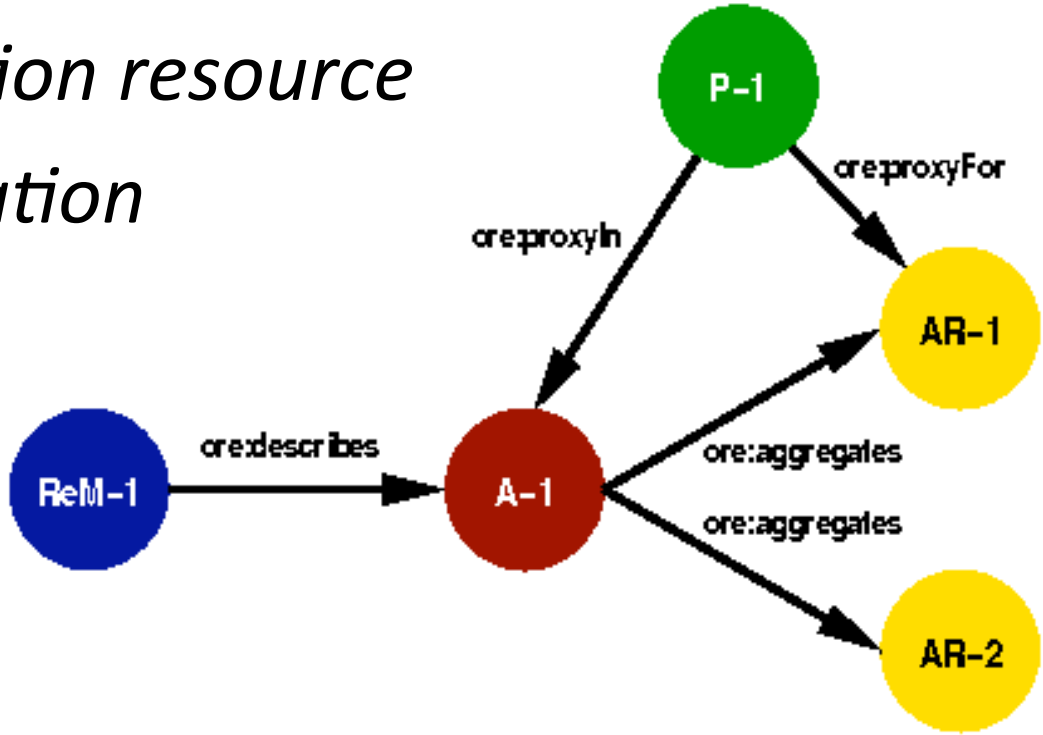
- Server supports 2 URIs for same XHTML+RDFa

```
S-1   = http://example.org/splash.html   (== foo.html)
A-1   = http://example.org/foo
ReM-1 = http://example.org/foo.html      (== splash.html)
ReM-2 = http://example.org/foo.atom
```




# Proxies

- Aggregated Resource *in context of Aggregation* (more later...)
- *Non-information resource*
  - *no representation*



# Requirements for Proxy URIs in HTTP

1. Redirect to the Aggregated Resource with HTTP status code "303 See Other" and

   Location: URI-AR

2. Indicate the Aggregation context with the Link header:

   Link: <URI-A>; rel="aggregation"

3. No restriction on URL syntax, but...




# ORE Proxy URI resolver

- Operated by OCLC (thanks!)
- Simple construction syntax:

http://oreproxy.org/r?what=URI-AR&where=URI-A

  - parameter order important
  - careful to URI encode (potentially doubly)
- Resolver has behavior just given
- Allows proxy URI use at no extra cost




```
simeon@rock ~>wget -S 'http://oreproxy.org/r?what=URI-AR&where=URI-A'

--23:19:56--  http://oreproxy.org/r?what=URI-AR&where=URI-A
           => `r?what=URI-AR&where=URI-A'
Resolving oreproxy.org... 132.174.1.110
Connecting to oreproxy.org|132.174.1.110|:80... connected.
HTTP request sent, awaiting response...
  HTTP/1.0 302 Found
  Location: http://oreproxy.org/oreproxy/r?what=URI-AR&where=URI-A
  Content-Length: 0
Location: http://oreproxy.org/oreproxy/r?what=URI-AR&where=URI-A [following]

--23:19:56--  http://oreproxy.org/oreproxy/r?what=URI-AR&where=URI-A
           => `r?what=URI-AR&where=URI-A'
HTTP request sent, awaiting response...
  HTTP/1.1 303 See Other
  Link: <URI-A>; rel="aggregation"
  Location: URI-AR
  Content-Type: text/plain;charset=ISO-8859-1
  Date: Mon, 16 Jun 2008 03:24:12 GMT
  Connection: close
Location: URI-AR [following]
--23:19:56--  http://oreproxy.org/oreproxy/URI-AR
           => `URI-AR'
```




# Resource Map Discovery

- Different Idioms for Design & Discovery

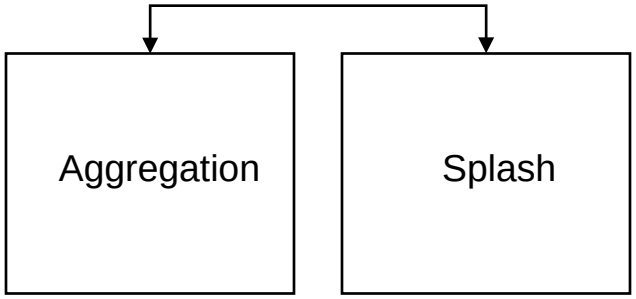
- Batch Discovery
  - OAI-PMH, SiteMaps, RSS/Atom
- Embedding Discovery Links
  - With HTML “link” element
  - With HTTP “Link” response header
- Exposing Aggregation URIs for copy-paste



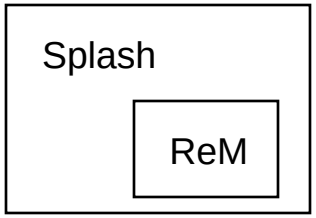


# Different Design Methods, Different Discovery Methods

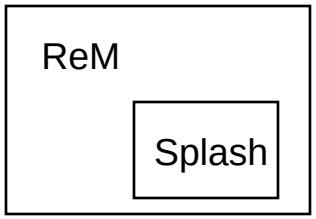
Syndication Format Idiom
(URI-A ≠ URI-S)



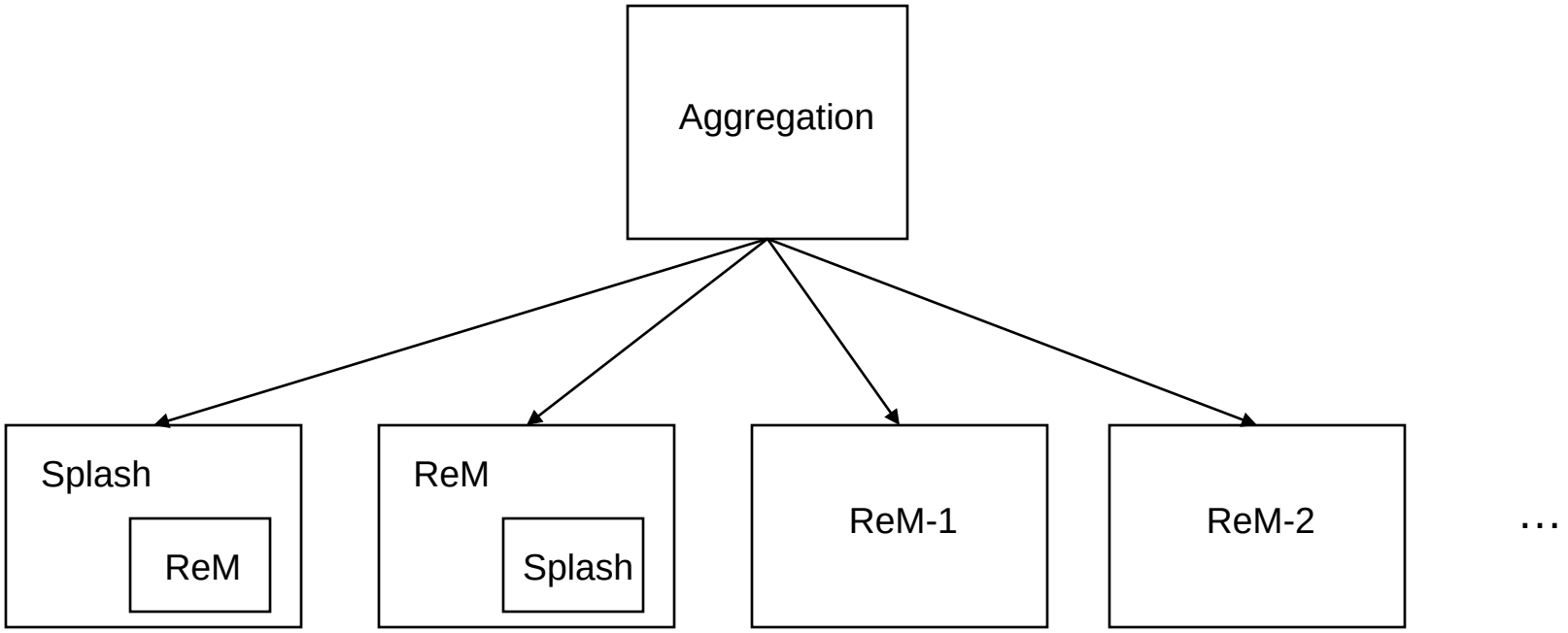

RDFa / Microformat Idiom
(URI-R = URI-S)




XML Stylesheet Idiom
(URI-R = URI-S)



# Putting it All Together…




Multiple values for URI-R and URI-S.
The only unique value is URI-A.

# Batch Discovery

- Resource Maps and Aggregations are resources and we already know how to expose large batches of resources:
  - OAI-PMH
  - SiteMaps
  - RSS/Atom




# Batch :: OAI-PMH

http://www.foo.edu/oai?verb=ListRecords&metadataPrefix=oai_rem_atom

```
<?xml version="1.0" encoding="UTF-8"?>
<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/"
         xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
         xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/
         http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
  <responseDate>2007-02-08T08:55:46Z</responseDate>
  <request verb=”ListRecords” metadataPrefix="oai_rem_atom">
        http://foo.edu/oai2</request>
  <ListRecords>
   <record>
    <header>
      <identifier>oai:foo.edu:object1</identifier>
      <datestamp>2007-01-06</datestamp>
    </header>
    <metadata>
        <!-- Insert object1 ReM here -->
    </metadata>
  </record>
  . . .
 </ListRecords>
</OAI-PMH>
```

MUST NOT
equal either ReM Atom /feed/id or /feed/link[@rel="self"]/@href

MUST be equal to ReM Atom /feed/updated

# OAI-PMH GetRecord Processing

http://www.foo.edu/oai?verb=GetRecord&identifier=oai:foo.edu:object1&metadataPrefix=oai_rem_atom

http://some.gateway.org/pmh2ore/www.foo.edu/oai?verb=GetRecord&metadataPefix=oai_rem_atom&identifier=oai:foo.edu:object1

```
<?xml version="1.0" encoding="UTF-8"?>
<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/"
         xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
         xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/
         http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
  <responseDate>2007-02-08T08:55:46Z</responseDate>
  <request verb="GetRecord" identifier="oai:foo.edu:object1"
           metadataPrefix="oai_rem_atom">http://foo.edu/oai2</request>
  <GetRecord>
   <record>
    <header>
      <identifier>oai:foo.edu:object1</identifier>
      <datestamp>2007-01-06</datestamp>
    </header>
    <metadata>
        <!-- Insert Object1 ReM here -->
    </metadata>
  </record>
 </GetRecord>
</OAI-PMH>
```

need a gateway to:

1. strip off OAI-PMH wrappers
2. return just what is inside <metadata>
3. reset the MIME type (e.g., from application/xml to application/atom+xml )

# Batch :: SiteMaps

http://www.foo.edu/sitemap-rem.xml

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
   <url>
      <loc>http://www.foo.edu/objects/object1.atom</loc>
      <lastmod>2007-01-06</lastmod>
   </url>
   <url>
      <loc>http://www.foo.edu/objects/object2.atom</loc>
      <lastmod>2007-08-11</lastmod>
      <changefreq>weekly</changefreq>
   </url>
   <url>
      <loc>http://www.foo.edu/objects/object3.atom</loc>
      <lastmod>2007-03-15T18:30:02Z</lastmod>
      <priority>0.3</priority>
   </url>
...
</urlset>
```

MUST equal /feed/link[@rel="self"]/@href or /feed/id for corresponding ReM

MUST be equal to ReM Atom /feed/updated




# Batch :: RSS

http://www.foo.edu/all-rems.rss

```
<?xml version="1.0"?>
<rss version="2.0">
  <channel>
    <title>ReMs at www.foo.edu</title>
    <link>http://www.foo.edu/</link>
    <description>All of the Resource Maps for resources at www.foo.edu</description>

    <item>
      <title>ReM for Object 1</title>
      <link>http://www.foo.org/objects/object1.atom</link>
      <description>ReM for Object 1</description>
      <pubDate>Sat, 06 Jan 2007 00:00:00 GMT</pubDate>
    </item>

    <item>
      <title>ReM for Object 2</title>
      <link>http://www.foo.org/objects/object2.atom</link>
      <description>ReM for Object 2</description>
      <pubDate>Sat, 11 Aug 2007 00:00:00 GMT</pubDate>
    </item>
</channel>
</rss>
```

MUST equal ReM Atom /feed/id;
or /feed/link[@rel="self"]/@href

MUST equal ReM Atom /feed/updated
(after conversion from RFC-822 format to ISO 8601 format)

# Batch :: Atom

http://www.foo.edu/all-rems.atom

```
<feed xmlns="http://www.w3.org/2005/Atom">
 <title>ReMs at www.foo.edu</title>
 <link href="http://www.foo.edu/" />
 <link href="http://www.foo.edu/all-rems.atom" rel="self"/>
 <updated>2007-08-15T18:30:02Z</updated>
 <author>
   <name>John Doe</name>
   <email>johndoe@foo.edu</email>
 </author>
 <id>urn:uuid:60a76c80-d399-11d9-b91C-0003939e0af6</id>

 <entry>
   <title>ReM For Object1</title>
   <link href="http://www.foo.org/objects/object1.atom"/>
   <id>urn:uuid:1225c695-cfb8-4ebb-aaaa-80da344efa6a</id>
   <updated>2007-01-06T00:00:00Z</updated>
 </entry>
 <entry>
   <title>ReM For Object2</title>
   <link href="http://www.foo.org/objects/object2.atom"/>
   <id>urn:uuid:9a2cc699-ccba-9e8b-132e-91da394e9a5c</id>
   <updated>2007-08-11T00:00:00Z</updated>
 </entry>
</feed>
```

MUST equal ReM Atom /feed/link[@rel="self"]/@href or /feed/id

MUST NOT equal ReM Atom /feed/id;

MUST equal ReM Atom /feed/updated

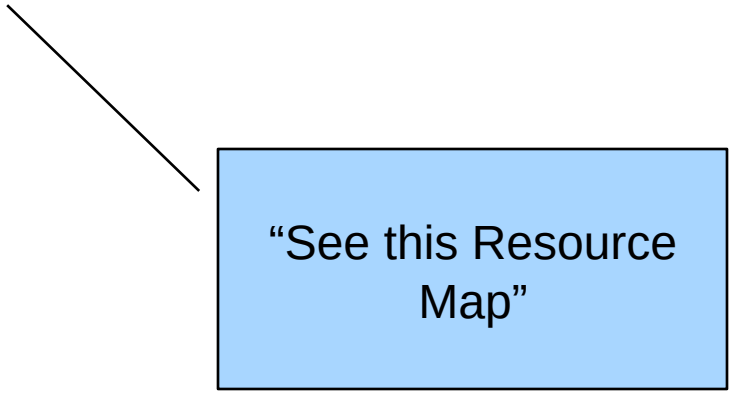
# Embedding Discovery Links into Resources

- Starting with a resource, how to find the associated Aggregations(s)?
  - HTML `<link>` element
  - HTTP Response Headers
  - Exposing URIs in HTML pages
- Depends on knowledge/control of Aggregated Resource by server/agent creating Aggregation




# HTML <link> element - 1

```
<html>
<head>
<title>Hello World.</title>
<link rel="resourcemap"
      href="http://example.net/hw.atom"
      type="application/atom+xml"/>
</head>
<body>
<img src="hello.jpeg"/>
<img src="world.jpeg"/>
</body>
</html>
```



# HTML <link> element - 2

```
<html>
<head>
<title>Hello World.</title>
<link rel="resourcemap alternate"
      href="http://example.net/hw.atom"
      type="application/atom+xml"/>
</head>
<body>
<img src="hello.jpeg"/>
<img src="world.jpeg"/>
</body>
</html>
```

Atom feed autodiscovery




# HTML <link> element - 3

```
<html>
<head>
<title>Hello World.</title>
<link rel="resourcemap alternate"
      href="http://example.net/hw.atom"
      type="application/atom+xml"/>
<link rel="resourcemap"
      href="http://example.net/hw.rdf"
      type="application/rdf+xml"/>
</head>
<body>
<img src="hello.jpeg"/>
<img src="world.jpeg"/>
</body>
</html>
```

RDF second Resource Map




# HTML <link> element - 4

```
<html>
<head>
<title>Hello World.</title>
<link rel="resourcemap alternate"
      href="http://example.net/hw.atom"
      type="application/atom+xml"/>
<link rel="resourcemap"
      href="http://example.net/hw.rdf"
      type="application/rdf+xml"/>
<link rel="resourcemap"
      href="http://example.net/hw"/>
</head>
<body>
<img src="hello.jpeg"/>
<img src="world.jpeg"/>
</body>
</html>
```

Link to Aggregation




# HTTP Response Headers

- For HTML and non-HTML resources

```
HEAD http://www.example.net/hello.jpeg HTTP/1.1
Host: www.example.net
Connection: close

HTTP/1.1 200 OK
Date: Sat, 26 May 2007 22:43:10 GMT
Server: Apache/2.2.0
Last-Modified: Sat, 26 May 2007 19:32:04 GMT
ETag: "c3596-816-92123500"
Accept-Ranges: bytes
Content-Length: 2070
Link: <http://example.net/hw.atom>; type="application/atom+xml";
  rel="resourcemap"
Content-Type: image/jpeg
Connection: close
```

Nottingham IETF Draft establishing semantic equivalence between HTML `<link>` and HTTP `Link:`

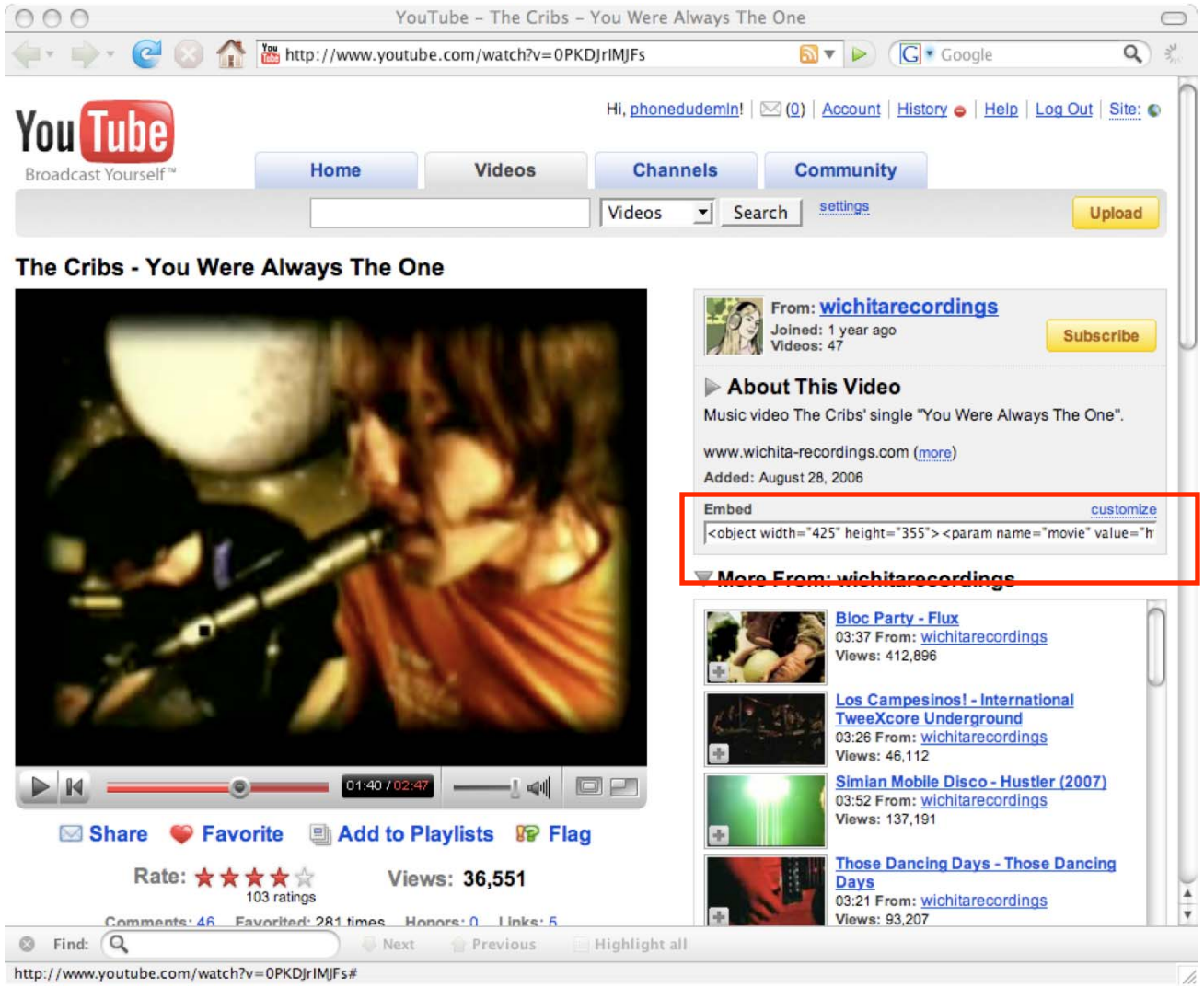



# Expose HTML and URI for copy-paste

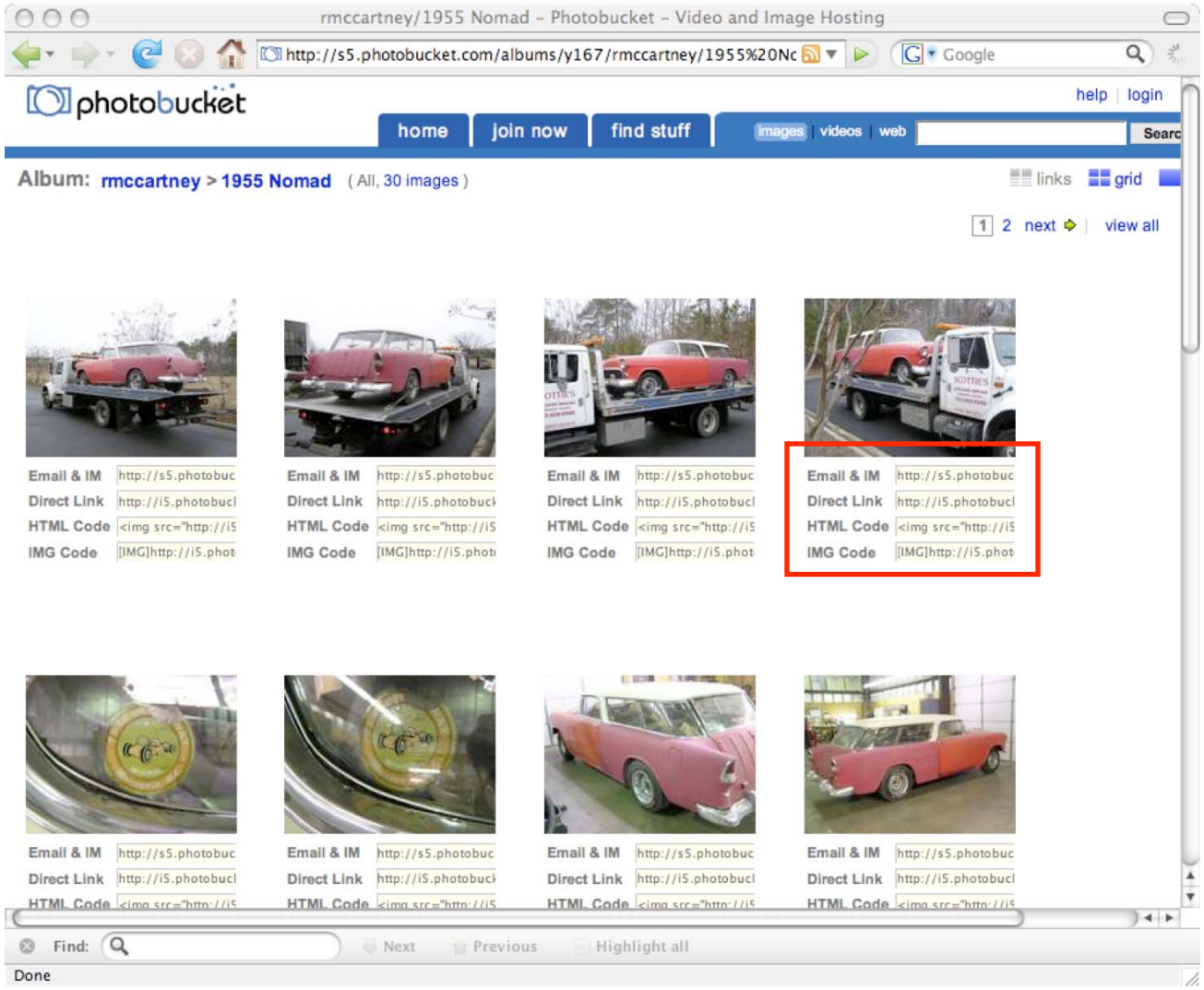
- There is precedent for exposing URIs, HTML, JavaScript, etc. as opaque strings for users to paste into other applications
- This is not the same as creating a hypertext link to the scripts…



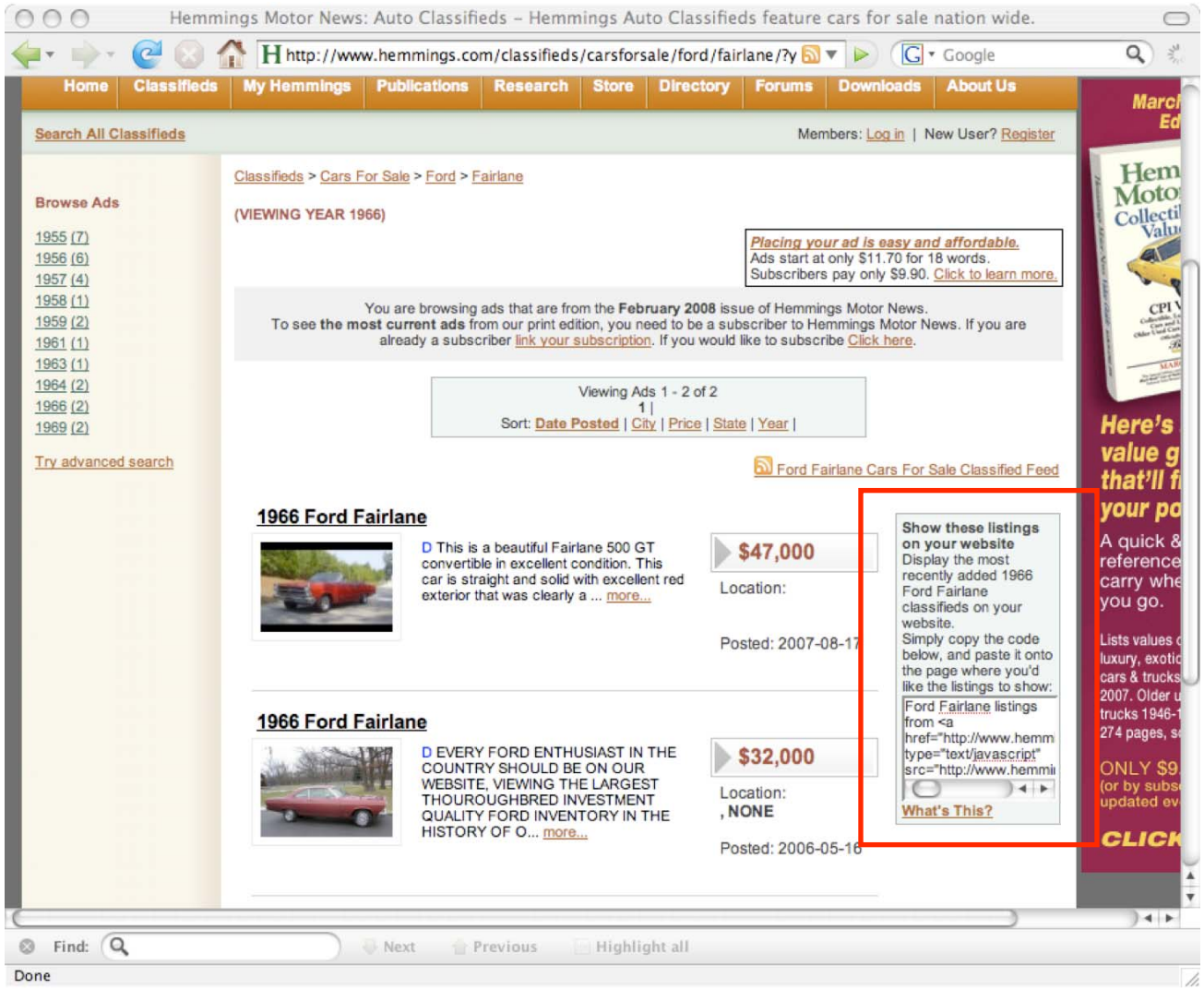

# HTML for copy-paste



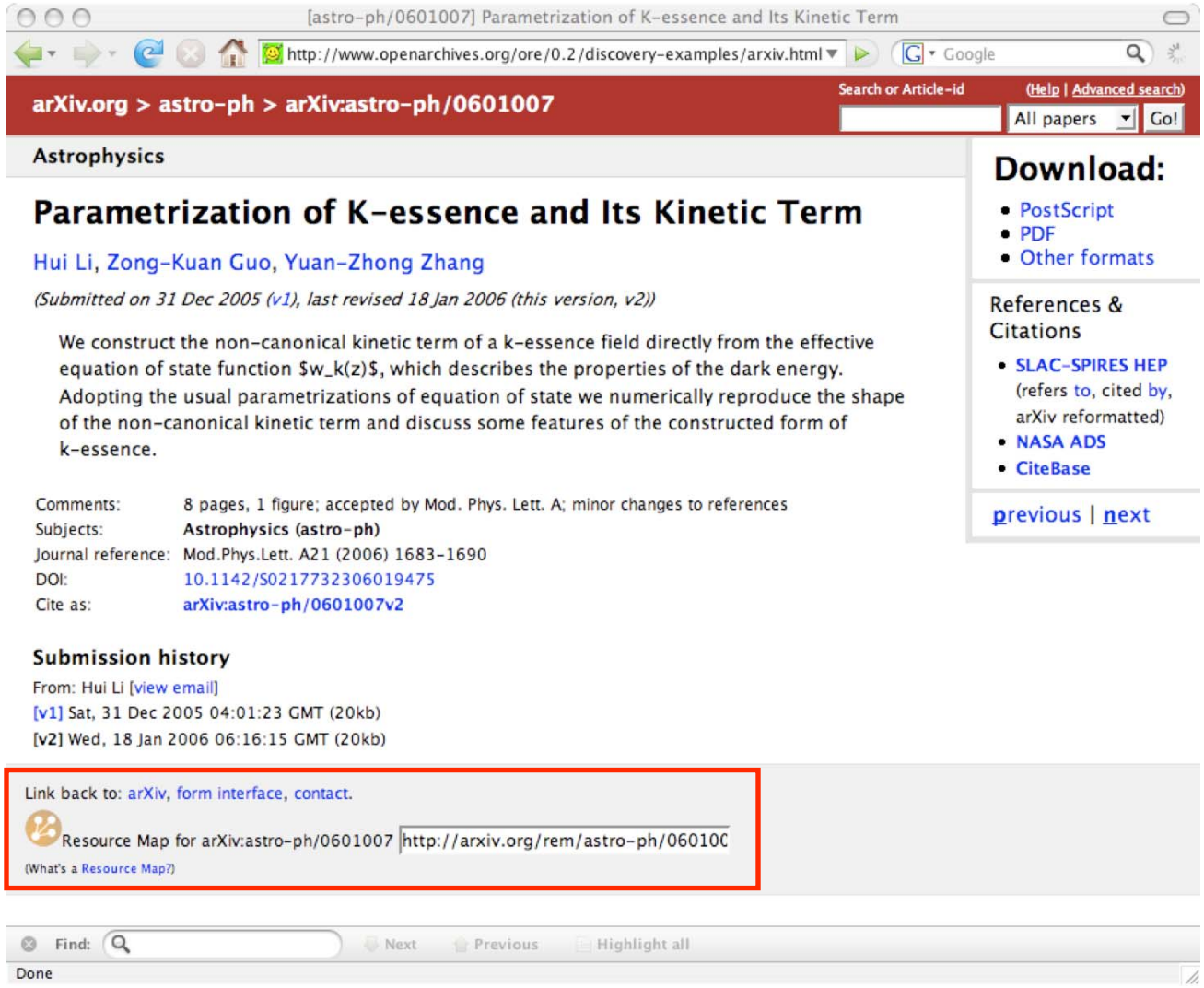
# URI & HTML for copy-paste



# HTML for copy-paste



# URI-R for copy-paste?



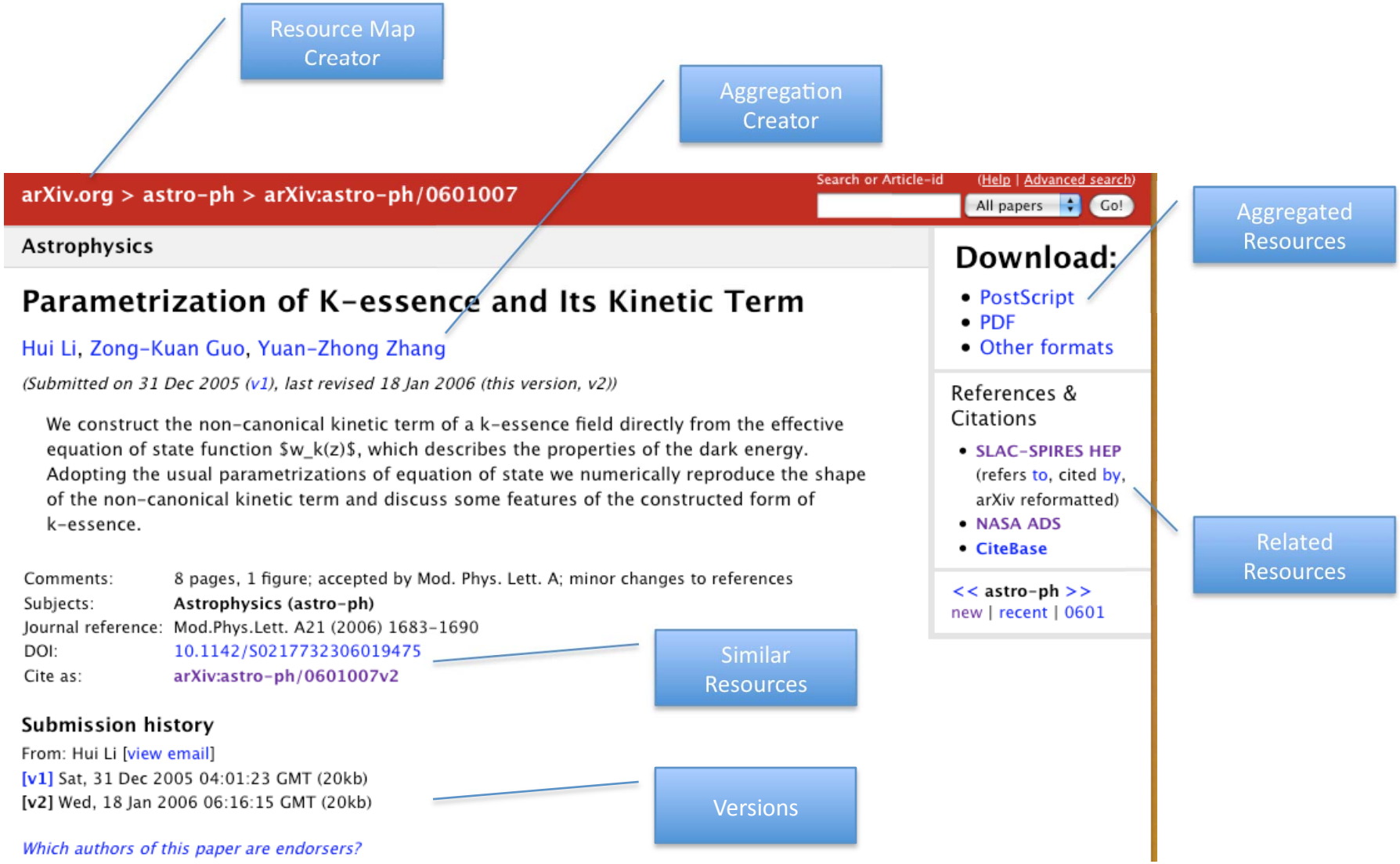
# Discovery is a Dirty Job

- Frequently a trade-off between “cleanliness” and “utility”
- Multiple discovery methods, possibly more evolving over time
- Each method has caveats and multiple opportunities to get it wrong

# Open Archives Initiative Object Reuse & Exchange

## Building a real example




http://arxiv.org/abs/astro-ph/0601007

# GRDDL

```
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns
```

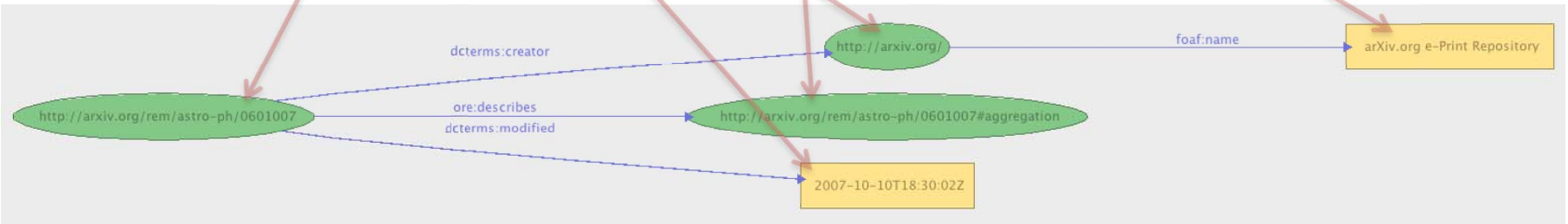
```
<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform" version="2.0"
    xmlns:xhtml="http://www.w3.org/1999/xhtml"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:atom="http://www.w3.org/2005/Atom"
    xmlns:ore="http://www.openarchives.org/ore/terms/"
    xmlns:dcterms="http://purl.org/dc/terms/"
    xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:foaf="http://xmlns.com/foaf/0.1/"
    xmlns:grddl="http://www.w3.org/2003/g/data-view#"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
    xmlns:xs="http://www.w3.org/2001/XMLSchema">

    <!-- atom-grddl.xsl  Version 0.9  May 2008 -->
    <!-- Crosswalk from ORE Atom serialization to RDF -->
    <!-- Los Alamos National Laboratary  -->
    <!-- Research Library -->
    <!-- Digital Library Research and Prototyping Team -->
    <!-- Author: Lyudmila Balakireva -->
    <!-- Email: ludab@lanl.gov -->
    <!-- [5 lines]
```

```
    <entry>

    <entry>

</feed>
```

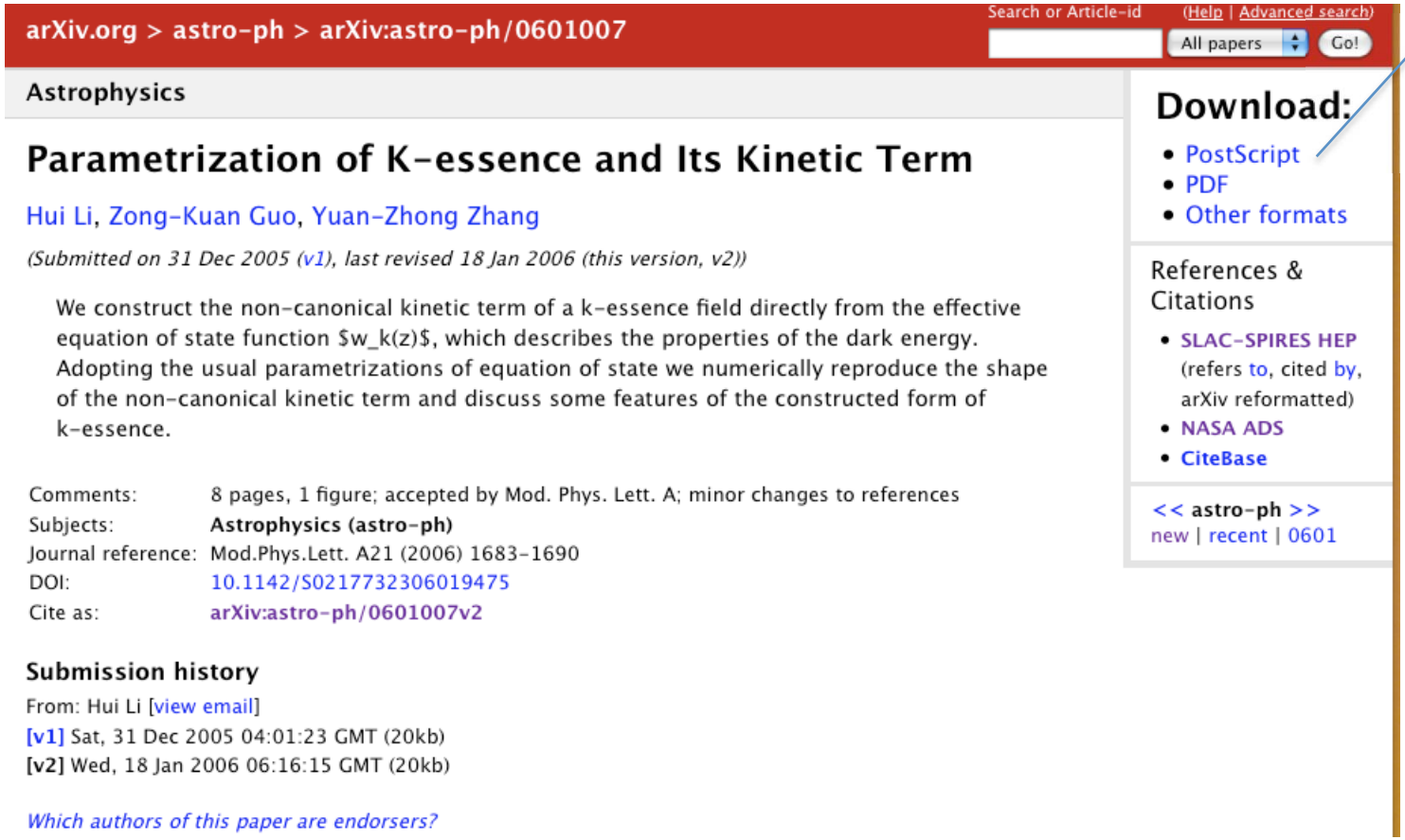
```
    <xsl:output method="xml" indent="yes" />

    <xsl:template match="atom:feed"> [167 lines]
```

# ReM Skeleton

```
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns="http://www.w3.org/2005/Atom">

    <id>http://arxiv.org/rem/astro-ph/0601007#aggregation</id>
    <link href="http://arxiv.org/rem/astro-ph/0601007" rel="self" type="application/atom+xml"/>
    <generator uri="http://arxiv.org/">arXiv.org e-Print Repository</generator>
    <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>
```



# Aggregated Resources




http://arxiv.org/abs/astro-ph/0601007

# Aggregated Resources

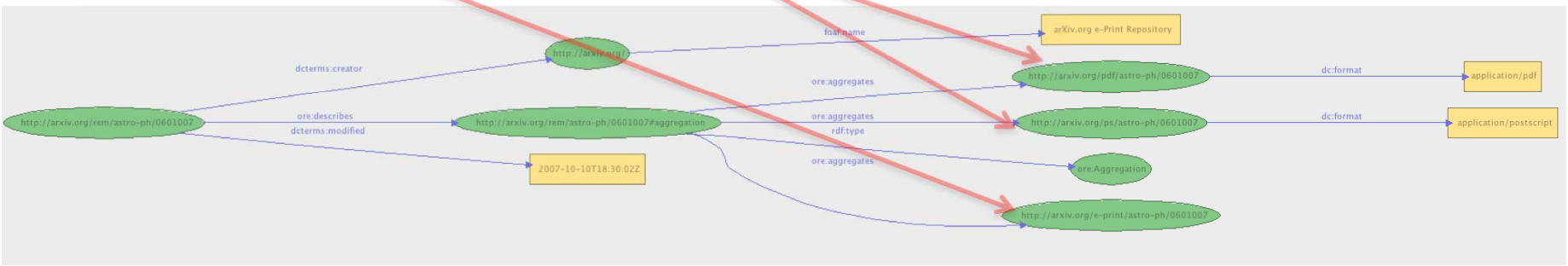
```
<feed xmlns="http://www.w3.org/2005/Atom">

    <id>http://arxiv.org/rem/astro-ph/0601007#aggregation</id>
    <link href="http://arxiv.org/rem/astro-ph/0601007" rel="self" type="application/atom+xml"/>
    <generator uri="http://arxiv.org/">arXiv.org e-Print Repository</generator>
    <updated>2007-10-10T18:30:02Z</updated>
    <category scheme="http://www.openarchives.org/ore/terms/" [1 line]

    <entry>
        <id>http://oreproxy.org/r?what=http://arxiv.org/ps/astro-ph/0601007&amp;where=http://arxiv.org/rem/astro-ph/0601007%23aggregation</id>
        <link href="http://arxiv.org/ps/astro-ph/0601007" rel="alternate"
            type="application/postscript"/>
    </entry>

    <entry>
        <id>http://oreproxy.org/r?what=http://arxiv.org/pdf/astro-ph/0601007&amp;where=http://arxiv.org/rem/astro-ph/0601007%23aggregation</id>
        <link href="http://arxiv.org/pdf/astro-ph/0601007" rel="alternate" type="application/pdf"/>
    </entry>

    <entry>
        <id>http://oreproxy.org/r?what=http://arxiv.org/e-print/astro-ph/0601007&amp;where=http://arxiv.org/rem/astro-ph/0601007%23aggregation</id>
        <link href="http://arxiv.org/e-print/astro-ph/0601007" rel="alternate"/>
    </entry>

</feed>
```

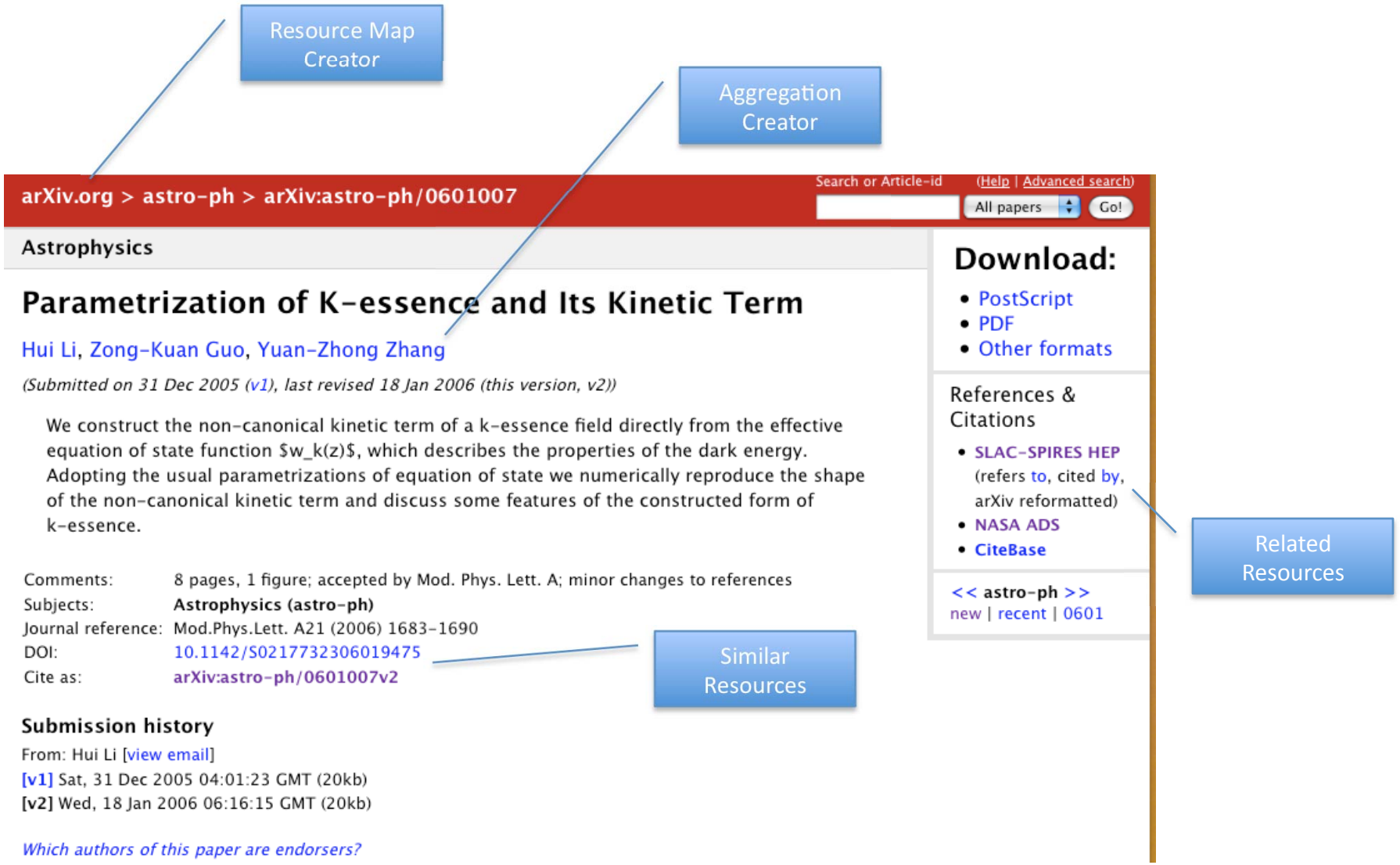
We'll talk about these later

ps version

Pdf version

oai-pmh metadata

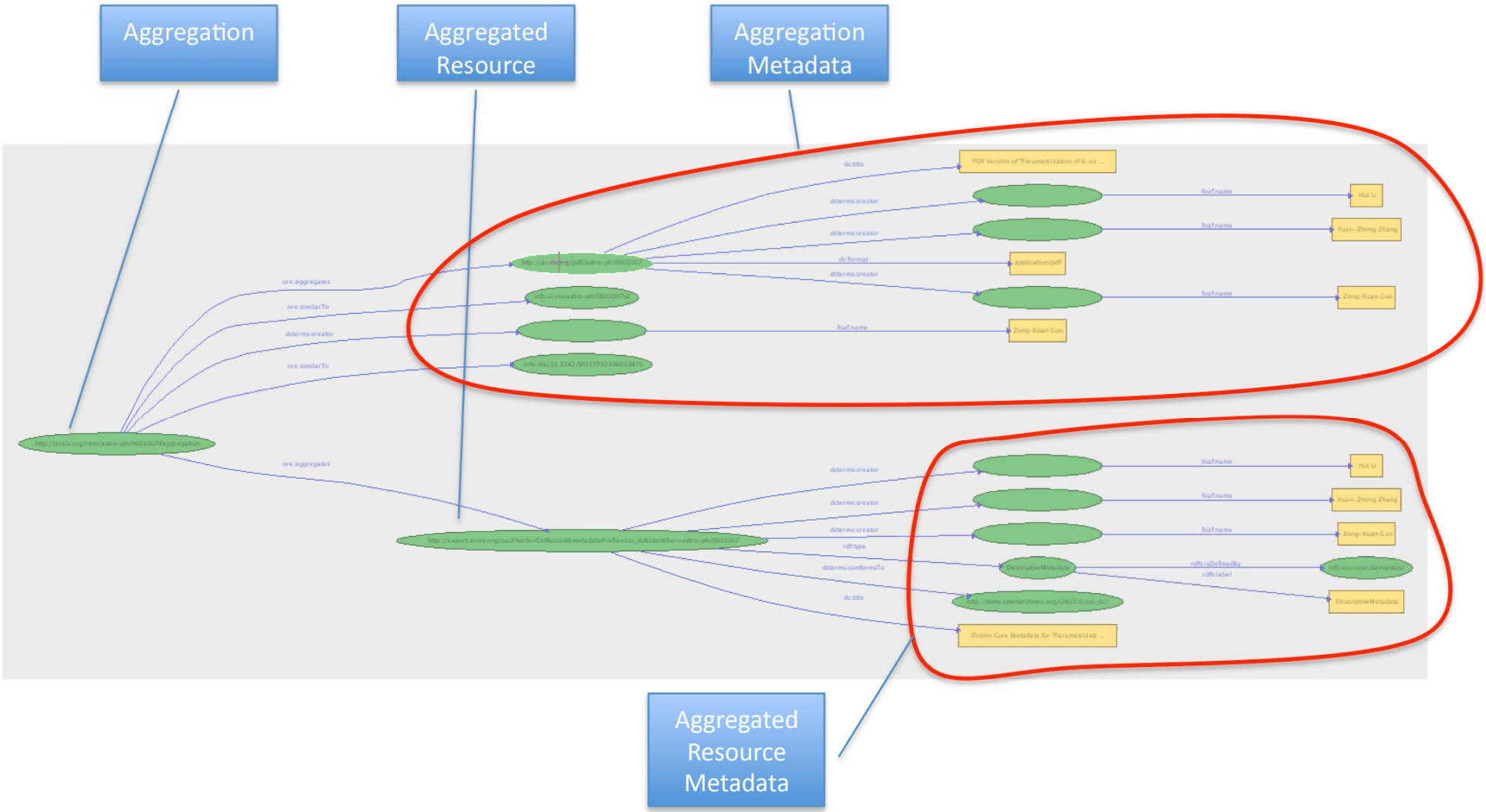


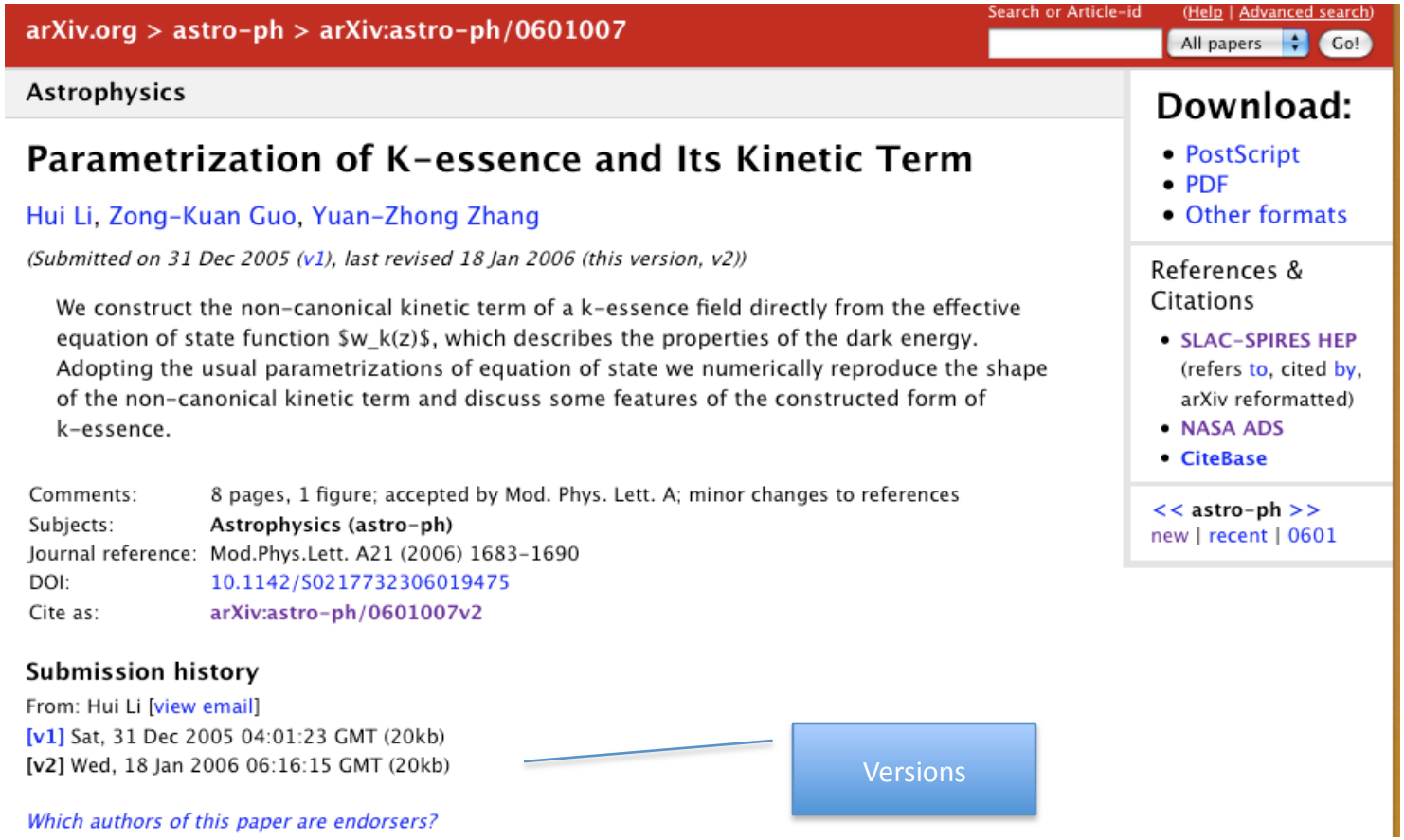



http://arxiv.org/abs/astro-ph/0601007

```
<feed xmlns="http://www.w3.org/2005/Atom"
      xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
      xmlns:dcterms="http://purl.org/dc/terms/">

    <id>http://arxiv.org/rem/astro-ph/0601007#aggregation</id>
    <link href="http://arxiv.org/rem/astro-ph/0601007" rel="self" type="application/atom+xml"/>
    <generator uri="http://arxiv.org/">arXiv.org e-Print Repository</generator>
    <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" />
    <link href="info:doi/10.1142/S0217732306019475" rel="related"/>
    <link href="info:arxiv/astro-ph/0601007v2" rel="related"/>
    <title>Parametrization of K-essence and Its Kinetic Term</title>
    <author><name>Hui Li</name></author>
    <author><name>Zong-Kuan Guo</name></author>
    <author><name>Yuan-Zhong Zhang</name></author>

    rdf:about="http://export.arxiv.org/oai2?verb=GetRecord&amp;metadataPrefix=oai_dc&amp;identifier=astro-ph/0601007">
    <dcterms:conformsTo rdf:resource="http://www.openarchives.org/OAI/2.0/oai_dc/"/>
</rdf:Description>

        <link href="http://arxiv.org/ps/astro-ph/0601007" rel="alternate" type="application/postscript"/>
        <title>PostScript Version of "Parametrization of K-essence and Its Kinetic Term"</title>
    </entry>

    <entry>
        <id>http://oreproxy.org/r?what=http://arxiv.org/pdf/astro-ph/0601007&amp;where=http://arxiv.org/rem/astro-ph/0601007%23aggregation</id>
        <link href="http://arxiv.org/pdf/astro-ph/0601007" rel="alternate" type="application/pdf"/>
        <title>PDF Version of "Parametrization of K-essence and Its Kinetic Term"</title>
    </entry>

    <entry>
        <id>http://oreproxy.org/r?what=http://arxiv.org/e-print/astro-ph/0601007&amp;where=http://arxiv.org/rem/astro-ph/0601007%23aggregation</id>
        <link href="http://arxiv.org/e-print/astro-ph/0601007" rel="alternate"/>
        <title>Other Versions of "Parametrization of K-essence and Its Kinetic Term"</title>
    </entry>
```





Aggregation
Aggregated Resource
Aggregation Metadata
ore:aggregates
ore:similarTo
dcterms:creator
ore:similarTo
ore:aggregates
dcterms:creator
dcterms:creator
dcterms:creator
rdf:type
dcterms:conformsTo
dc:title
dc:title
dcterms:creator
dcterms:creator
dcterms:creator
dc:format
dcterms:creator
foaf:name
foaf:name
foaf:name
foaf:name
foaf:name
Hui Li
Yuan-Zhong Zhang
Zong-Kuan Guo
Zong-Kuan Guo
Hui Li
Yuan-Zhong Zhang
Zong-Kuan Guo
rdfs:isDefinedBy
rdfs:label
Aggregated Resource Metadata




http://arxiv.org/abs/astro-ph/0601007

```
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns="http://www.w3.org/2005/Atom"
    xmlns:grddl="http://www.w3.org/2003/g/data-view#"
    grddl:transformation="http://www.openarchives.org/ore/atom-grddl.xsl"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:dcterms="http://purl.org/dc/terms/">

    <id>http://arxiv.org/rem/astro-ph/0601007#aggregation</id>
    <link href="http://arxiv.org/rem/astro-ph/0601007" rel="self" type="application/atom+xml"/>
    <generator uri="http://arxiv.org/">arXiv.org e-Print Repository</generator>
    <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" />
    <link href="info:doi/10.1142/S0217732306019475" rel="related"/>
    <link href="info:arxiv/astro-ph/0601007v2" rel="related"/>
    <link href="http://jp.arxiv.org/rem/astro-ph/0601007#aggregation" rel="related"/>
    <rdf:Description rdf:about="http://arxiv.org/rem/astro-ph/0601007">
        <dcterms:hasVersion rdf:resource="http://arxiv.org/rem/astro-ph/0601007v1"/>
    </rdf:Description>
    <title>Parametrization of K-essence and its Kinetic Term</title>
    <author><name>Hui Li</name></author>
    <author><name>Zong-Kuan Guo</name></author>
    <author><name>Yuan-Zhong Zhang</name></author>

    <entry> [12 lines]

    <entry> [8 lines]

    <entry> [8 lines]

    <entry> [5 lines]

</feed>
```




So what are these funny URL's

```
<feed xmlns="http://www.w3.org/2005/Atom">

    <id>http://arxiv.org/rem/astro-ph/0601007#aggregation</id>
    <link href="http://arxiv.org/rem/astro-ph/0601007" rel="self" type="application/atom+xml"/>
    <generator uri="http://arxiv.org/">arXiv.org e-Print Repository</generator>
    <updated>2007-10-10T18:30:02Z</updated>
    <category scheme="http://www.openarchives.org/ore/terms/" [1 line]

    <entry>
        <id>http://oreproxy.org/r?what=http://arxiv.org/ps/astro-ph/0601007&amp;where=http://arxiv.org/rem/astro-ph/0601007%23aggregation</id>
        <link href="http://arxiv.org/ps/astro-ph/0601007" rel="alternate"
            type="application/postscript"/>
    </entry>

    <entry>
        <id>http://oreproxy.org/r?what=http://arxiv.org/pdf/astro-ph/0601007&amp;where=http://arxiv.org/rem/astro-ph/0601007%23aggregation</id>
        <link href="http://arxiv.org/pdf/astro-ph/0601007" rel="alternate" type="application/pdf"/>
    </entry>

    <entry>
        <id>http://oreproxy.org/r?what=http://arxiv.org/e-print/astro-ph/0601007&amp;where=http://arxiv.org/rem/astro-ph/0601007%23aggregation</id>
        <link href="http://arxiv.org/e-print/astro-ph/0601007" rel="alternate"/>
    </entry>

</feed>
```

# Atom Entry IDs

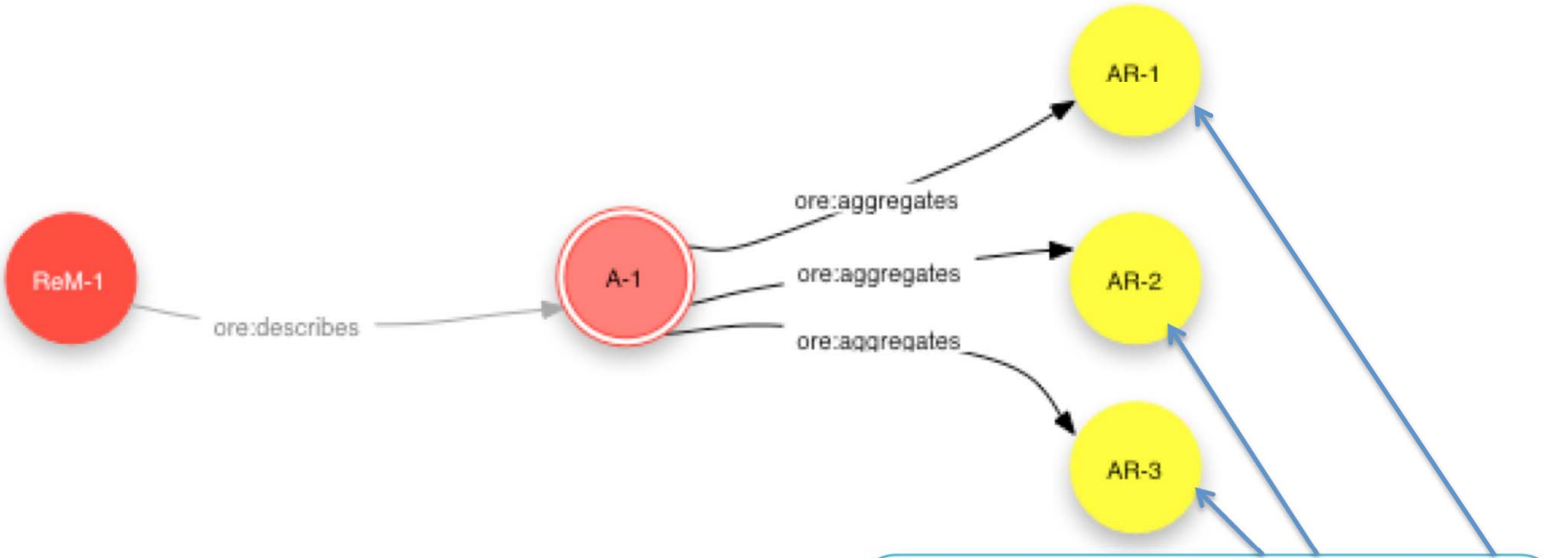
```
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns="http://www.w3.org/2005/Atom">

    <title>Example Feed</title>
    <link href="http://example.org/"/>
    <updated>2003-12-13T18:30:02Z</updated>
    <author>
        <name>John Doe</name>
    </author>
    <id>urn:uuid:60a76c80-d399-11d9-b93C-0003939e0af6</id>

    <entry>
        <title>Atom-Powered Robots Run Amok</title>
        <link href="http://example.org/2003/12/13/atom03"/>
        <id>urn:uuid:1225c695-cfb8-4ebb-aaaa-80da344efa6a</id>
        <updated>2003-12-13T18:30:02Z</updated>
        <summary>Some text.</summary>
    </entry>

    <entry>
        <title>Atom-Powered Spaceships Battle Robots</title>
        <link href="http://example.org/2003/12/13/atom04"/>
        <id>urn:uuid:1225c695-cfb8-4ebb-aaaa-80da344abf7f</id>
        <updated>2003-12-13T18:30:02Z</updated>
        <summary>Some other text.</summary>
    </entry>

</feed>
```

- Globally unique
  - urn
  - tag uri
    - Internal key
    - Timestamp
- "travels" with entry

We have to return to the data model for a little context…..

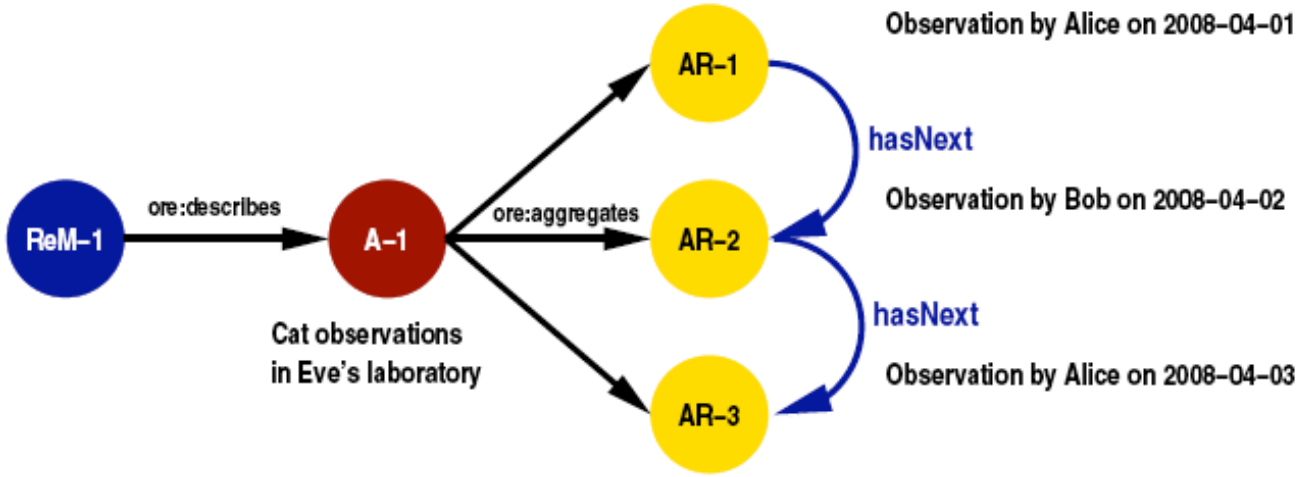



# Basic ORE Model

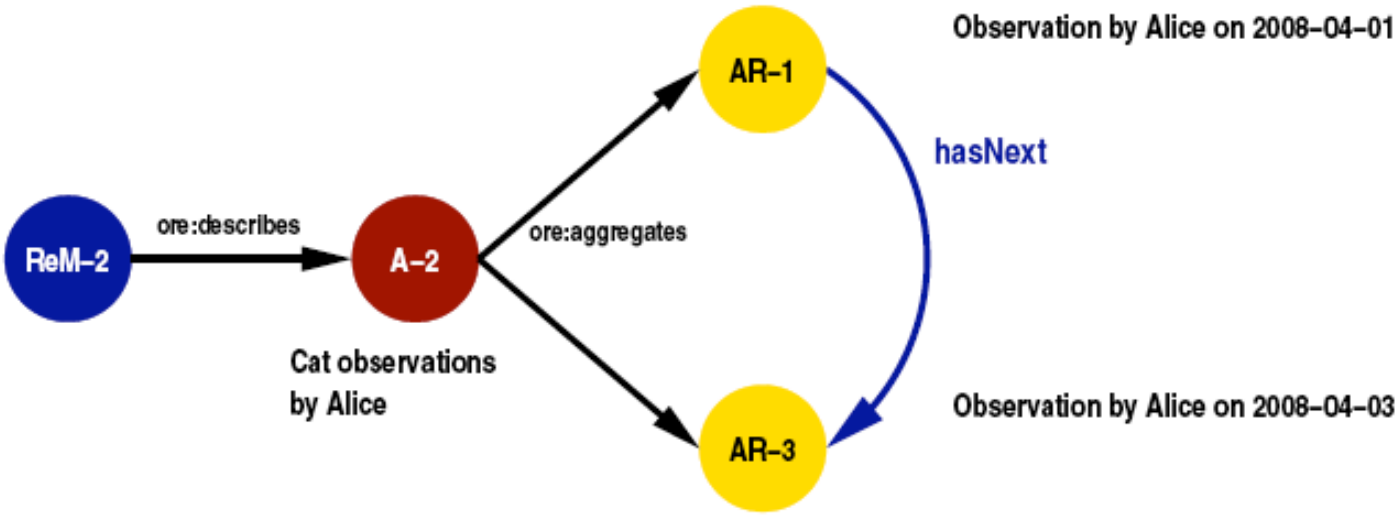



- These URIs are not special to the aggregation
- Example: paper on my web site that is a member of Simeon's favorite paper" collection

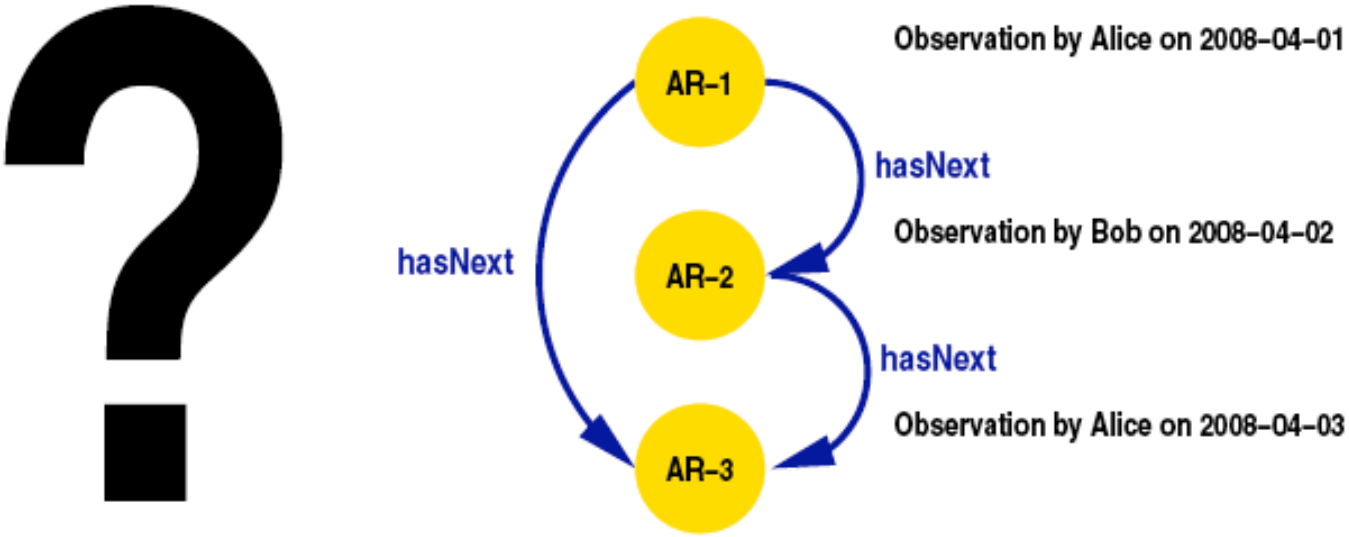
# Alice and Bob observe cats in Eve's laboratory



# 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*

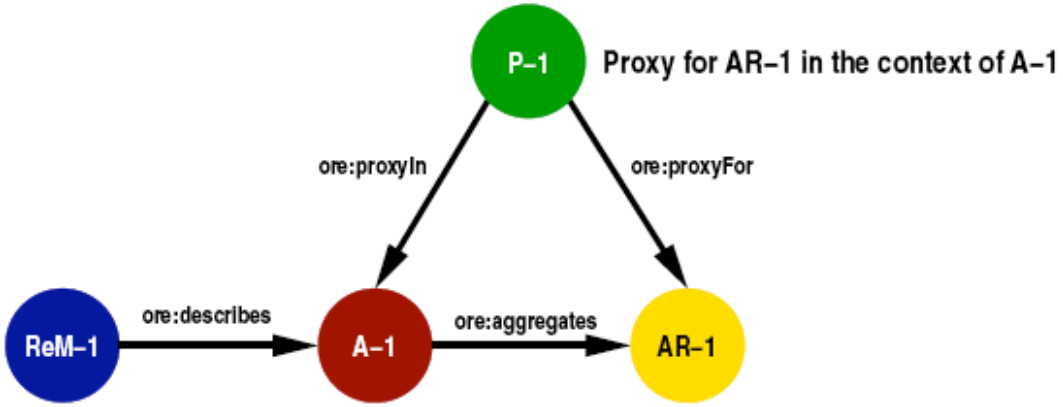



## Modelling a resource in context

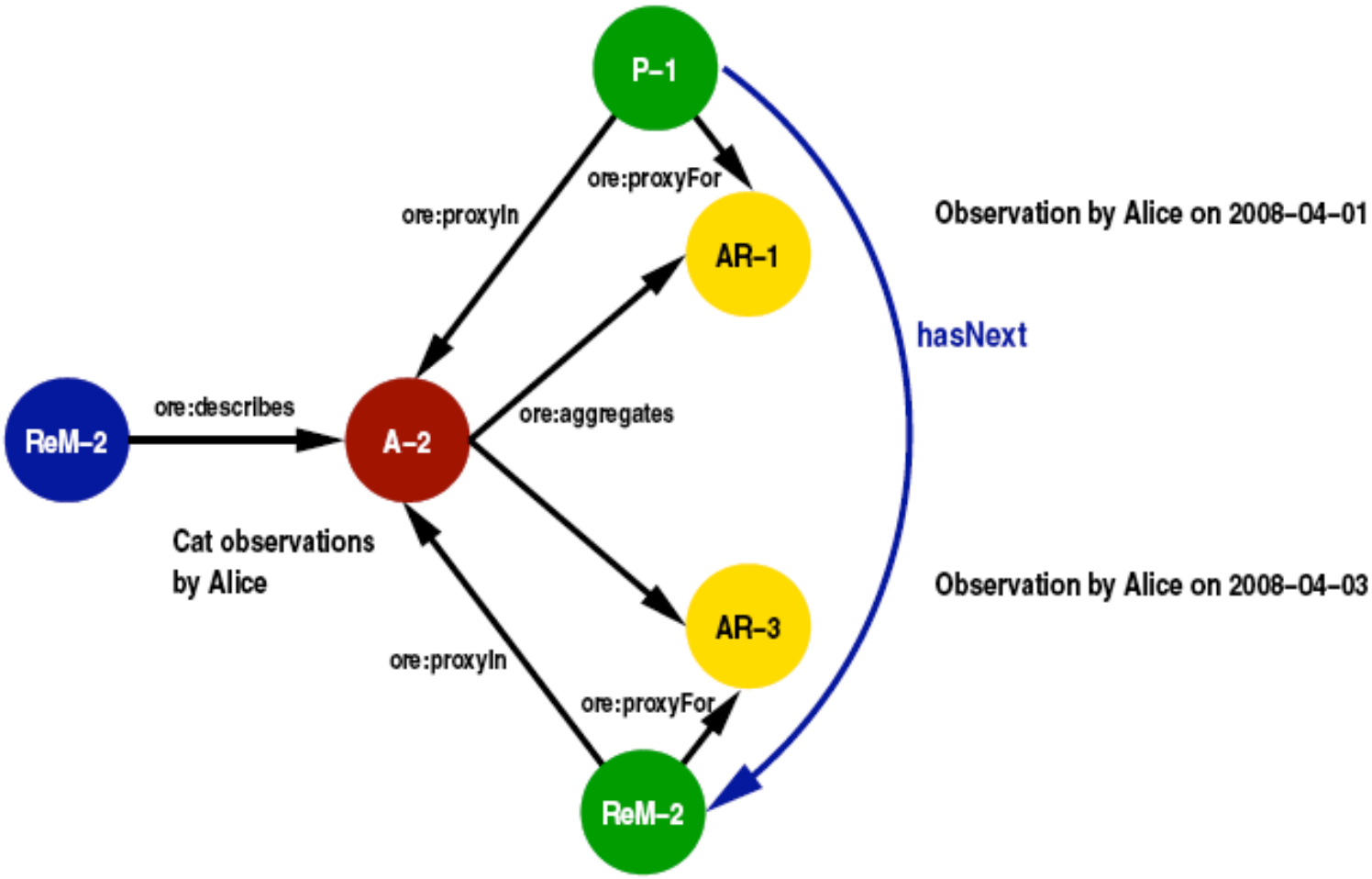
Two components:

- 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



# Rules for Proxy URIs

- Must be unique to a specific Aggregation (URI -A) and a specific Aggregated Resource (URI -AR)
    - "stand for" URI-AR in context of URI-A
- Resolution of Proxy URI MUST
  - Redirect client to URI-AR (HTTP 303 – see other)
  - Indicate URI-A in link header

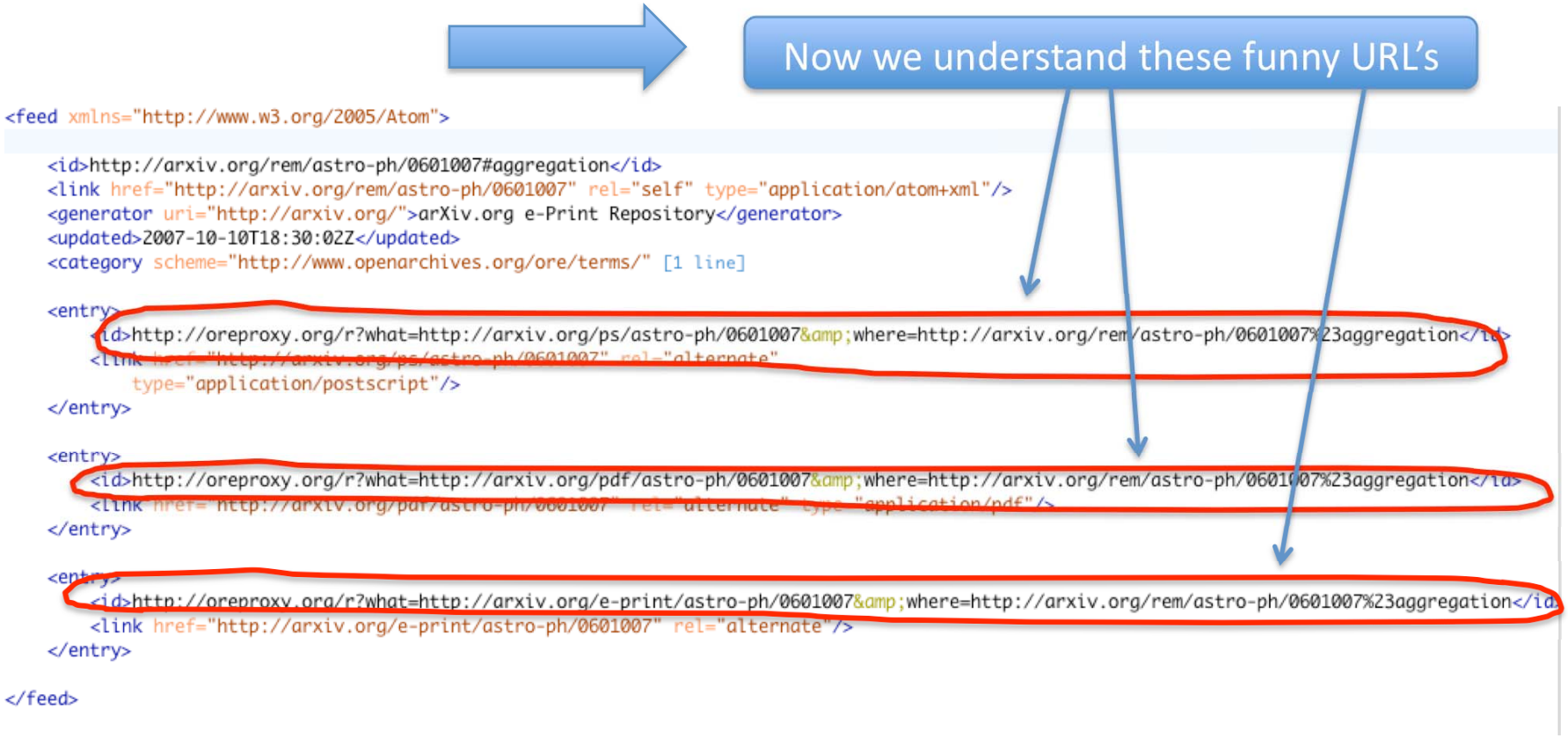
```
Link: <URI-A>; rel="aggregation"
```




# ORE Proxy URI resolver

- At http://oreproxy/r (supported by OCLC)
- Standard syntax
  - http://oreproxy.org/r?what=URI-AR&where=URI-A
  - URI-AR and URI-A MUST be encoded





Now we understand these funny URL's
<feed xmlns="http://www.w3.org/2005/Atom">
<id>http://arxiv.org/rem/astro-ph/0601007#aggregation</id>
<link href="http://arxiv.org/rem/astro-ph/0601007" rel="self" type="application/atom+xml"/>
<generator uri="http://arxiv.org/">arXiv.org e-Print Repository</generator>
<updated>2007-10-10T18:30:02Z</updated>
<category scheme="http://www.openarchives.org/ore/terms/" [1 line]
<entry>
<id>http://oreproxy.org/r?what=http://arxiv.org/ps/astro-ph/0601007&where=http://arxiv.org/rem/astro-ph/0601007%23aggregation</id>
type="application/postscript"/>
</entry>
<entry>
<id>http://oreproxy.org/r?what=http://arxiv.org/pdf/astro-ph/0601007&where=http://arxiv.org/rem/astro-ph/0601007%23aggregation</id>
</entry>
<entry>
<id>http://oreproxy.org/r?what=http://arxiv.org/e-print/astro-ph/0601007&where=http://arxiv.org/rem/astro-ph/0601007%23aggregation</id>
<link href="http://arxiv.org/e-print/astro-ph/0601007" rel="alternate"/>
</entry>
</feed>

# Lineage: Proxies in the Wild

- A Resource may be in multiple Aggregations
- How can we indicate provenance
    - "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.

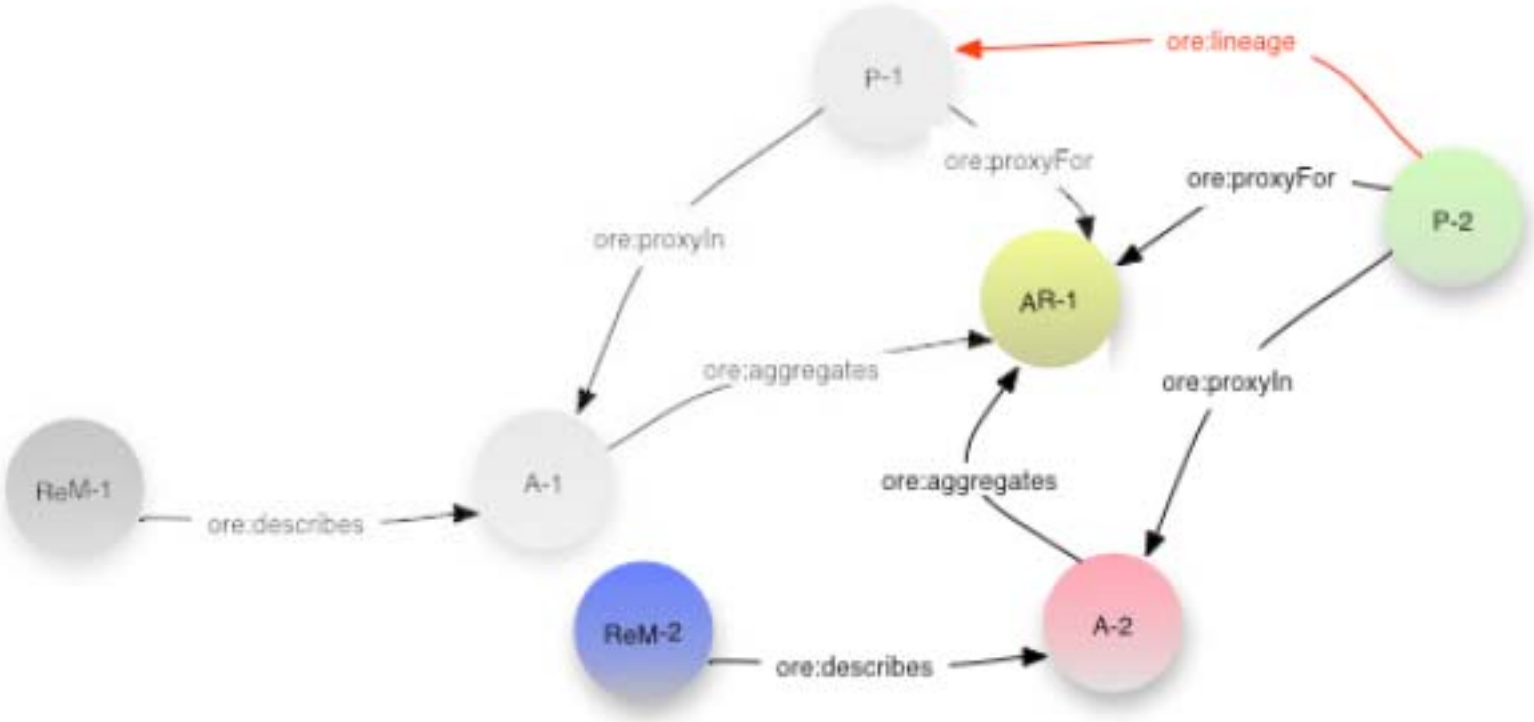
$\Longrightarrow$ use `ore:lineage` as relation between proxy nodes





```
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns="http://www.w3.org/2005/Atom"
    xmlns:grddl="http://www.w3.org/2003/g/data-view#"
    grddl:transformation="http://www.openarchives.org/ore/atom-grddl.xsl"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:dcterms="http://purl.org/dc/terms/">

    <id>http://converter.repo.org/converted/2008-03-30/#aggregation</id>
    <link href="http://converter.repo.org/converted/2008-03-30/" rel="self" type="application/atom+xml"/>
    <generator uri="http://repo.org/">OpenURL PDF Converter Repository</generator>
    <updated>2008-03-30T13:03:20Z</updated>
    <category scheme="http://www.openarchives.org/ore/terms/"
        term="http://www.openarchives.org/ore/terms/Aggregation" label="Aggregation" />

    <!-- other entries -->

    <entry>
        <id>http://oreproxy.org/r?
            what=http://repo.org/astro-ph-0601007.pdf&amp;where=http://converter.repo.org/converted/2008-03-30/%23aggregation</id>
        <link href="http://repo.org/astro-ph-0601007.pdf" rel="alternate"/>
        <updated>2008-03-30T13:03:20Z</updated>
        <title>"Parametrization of K-essence and Its Kinetic Term" (with OpenURL-enabled references)</title>
        <link href="http://oreproxy.org/r?
            what=http://arxiv.org/e-print/astro-ph/0601007&amp;where=http://arxiv.org/rem/astro-ph/0601007%23aggregation" rel="via"/>
    </entry>

    <!-- other entries -->

</feed>
```





ore:lineage
P-1
ore:proxyFor
ore:proxyFor
P-2
ore:proxyIn
AR-1
ore:aggregates
ore:proxyIn
ReM-1
ore:describes
A-1
ore:aggregates
ReM-2
ore:describes
A-2

# Open Archives Initiative Object Reuse & Exchange

Questions? Feedback?