

# Open Archives Initiative Object Re-Use & Exchange

Carl Lagoze <sup>(1)</sup> , Herbert Van de Sompel <sup>(2)</sup>

<sup>(1)</sup> Information Science, Cornell University  
lagoze@cs.cornell.edu  
<http://www.cs.cornell.edu/lagoze>

<sup>(2)</sup> Digital Library Research & Prototyping Team, Research Library, Los Alamos National Laboratory  
herbertv@lanl.gov  
<http://public.lanl.gov/herbertv>



ORE is supported by the Andrew W. Mellon Foundation  
with additional support of the National Science Foundation and Microsoft



# OAI Object Re-Use and Exchange

- OAI-ORE is a new interoperability effort conducted under the umbrella of the OAI
- Supported by the **Andrew W. Mellon Foundation**; additional support from the **National Science Foundation** and **Microsoft**
- International effort; October 2006 - September 2008:
  - Coordinators: Carl Lagoze & Herbert Van de Sompel
  - ORE Technical Committee: 13 international members
  - ORE Liaison Group: 8 international members
  - ORE Advisory Committee: 16 international members
  - Representing: scholarly publishers and aggregators, eScience, eHumanities, education, search engines, various repository systems, digital library efforts, related standardization efforts, etc.
- See <http://www.openarchives.org/ore/>
- See <http://www.openarchives.org/ore/documents/CompoundObjects-200705.html> for a recent white paper



# OAI Object Re-Use and Exchange

## **Core goal of OAI-ORE:**

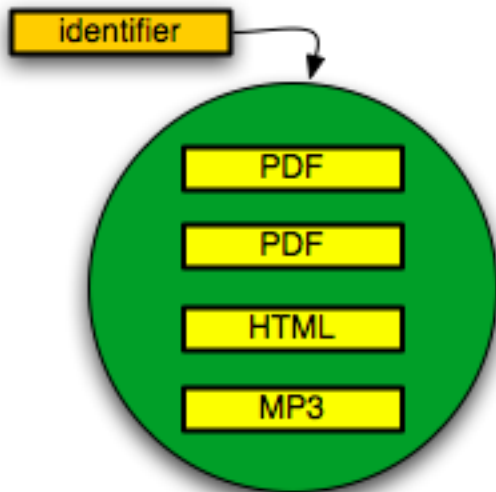
**Facilitate Use and Re-Use of Compound Information Objects (and of their component parts)**



# Compound Information Objects

Units of scholarly communication are compound information objects:

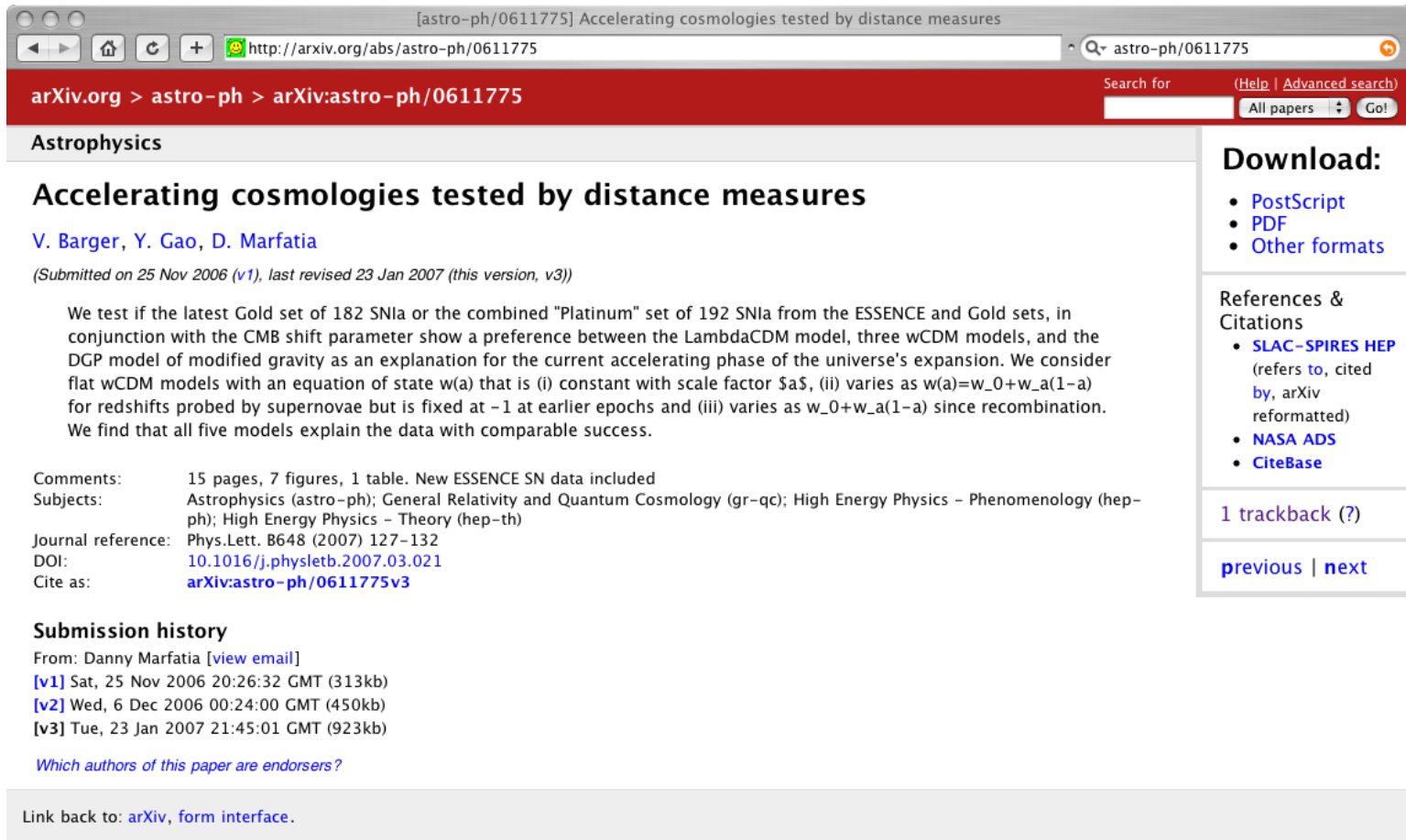
Identified, bounded aggregations of related information units that form a logical whole.



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

# Scholarly Examples



[astro-ph/0611775] Accelerating cosmologies tested by distance measures

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

arXiv.org > astro-ph > arXiv:astro-ph/0611775

Astrophysics

## Accelerating cosmologies tested by distance measures

V. Barger, Y. Gao, D. Marfatia

(Submitted on 25 Nov 2006 (v1), last revised 23 Jan 2007 (this version, v3))

We test if the latest Gold set of 182 SNIa or the combined "Platinum" set of 192 SNIa from the ESSENCE and Gold sets, in conjunction with the CMB shift parameter show a preference between the LambdaCDM model, three wCDM models, and the DGP model of modified gravity as an explanation for the current accelerating phase of the universe's expansion. We consider flat wCDM models with an equation of state  $w(a)$  that is (i) constant with scale factor  $a$ , (ii) varies as  $w(a)=w_0+w_a(1-a)$  for redshifts probed by supernovae but is fixed at  $-1$  at earlier epochs and (iii) varies as  $w_0+w_a(1-a)$  since recombination. We find that all five models explain the data with comparable success.

Comments: 15 pages, 7 figures, 1 table. New ESSENCE SN data included  
Subjects: Astrophysics (astro-ph); General Relativity and Quantum Cosmology (gr-qc); High Energy Physics - Phenomenology (hep-ph); High Energy Physics - Theory (hep-th)  
Journal reference: Phys.Lett. B648 (2007) 127-132  
DOI: 10.1016/j.physletb.2007.03.021  
Cite as: arXiv:astro-ph/0611775v3

### Submission history

From: Danny Marfatia [view email]  
[v1] Sat, 25 Nov 2006 20:26:32 GMT (313kb)  
[v2] Wed, 6 Dec 2006 00:24:00 GMT (450kb)  
[v3] Tue, 23 Jan 2007 21:45:01 GMT (923kb)

*Which authors of this paper are endorsers?*

Link back to: arXiv, form interface.

**Download:**

- PostScript
- PDF
- Other formats

**References & Citations**

- SLAC-SPIRES HEP (refers to, cited by, arXiv reformatted)
- NASA ADS
- CiteBase

1 [trackback \(?\)](#)

[previous](#) | [next](#)

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

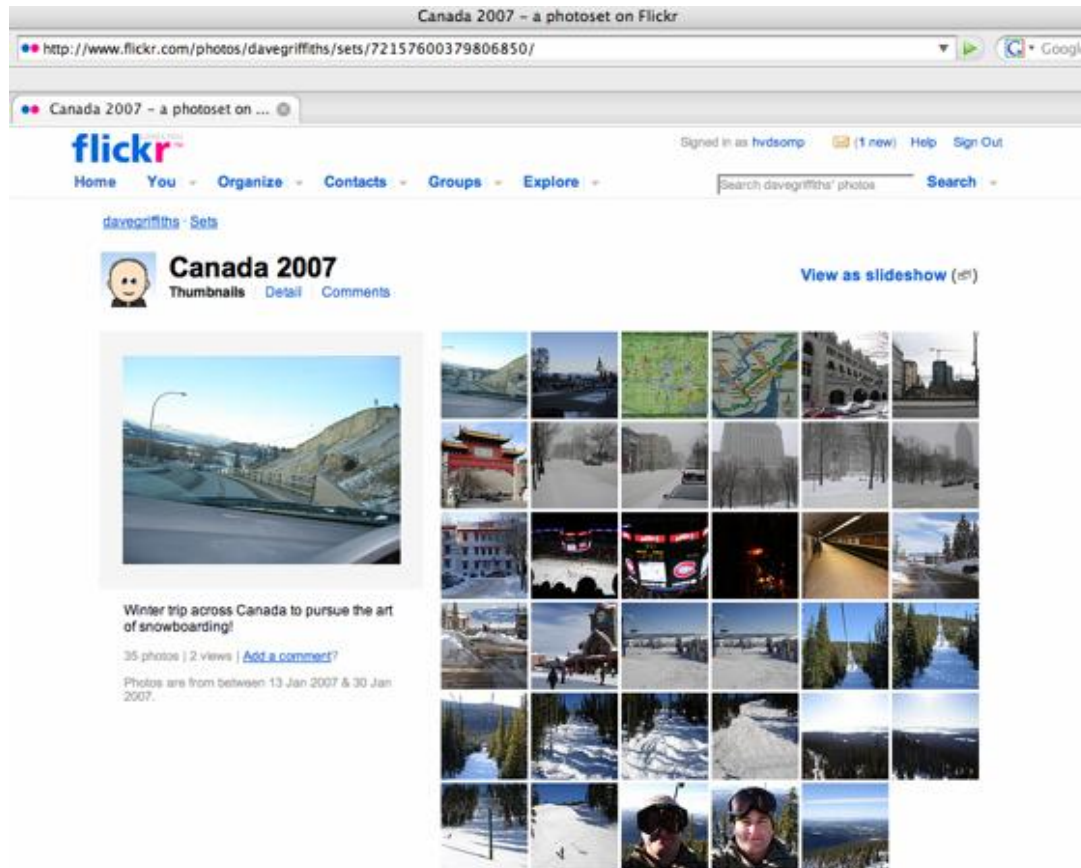


## And more scholarly examples ...

- Scholarly publication with an article and supporting information including dataset, video, etc.
- Digitized book with multiple chapters, each chapter containing multiple scanned pages.
- Archaeological assemblies of images, maps, charts, and find lists.
- An ARTstor image object that is the aggregation of various renderings of the same source image.
- ...



But these things are not only scholarly ...



The screenshot shows a web browser window displaying a Flickr photoselection. The browser's address bar shows the URL: <http://www.flickr.com/photos/davegriffiths/sets/72157600379806850/>. The Flickr page header includes the site logo, navigation links (Home, You, Organize, Contacts, Groups, Explore), and a search bar. The main content area features the title "Canada 2007" with a "View as slideshow" link. Below the title is a grid of 35 photo thumbnails. A larger thumbnail on the left shows a snowy landscape with a road and a hill. Below this larger thumbnail is the text: "Winter trip across Canada to pursue the art of snowboarding!" followed by "35 photos | 2 views | [Add a comment?](#)" and "Photos are from between 13 Jan 2007 & 30 Jan 2007."

<http://www.flickr.com/photos/davegriffiths/sets/72157600379806850/>



# OAI Object Re-Use and Exchange

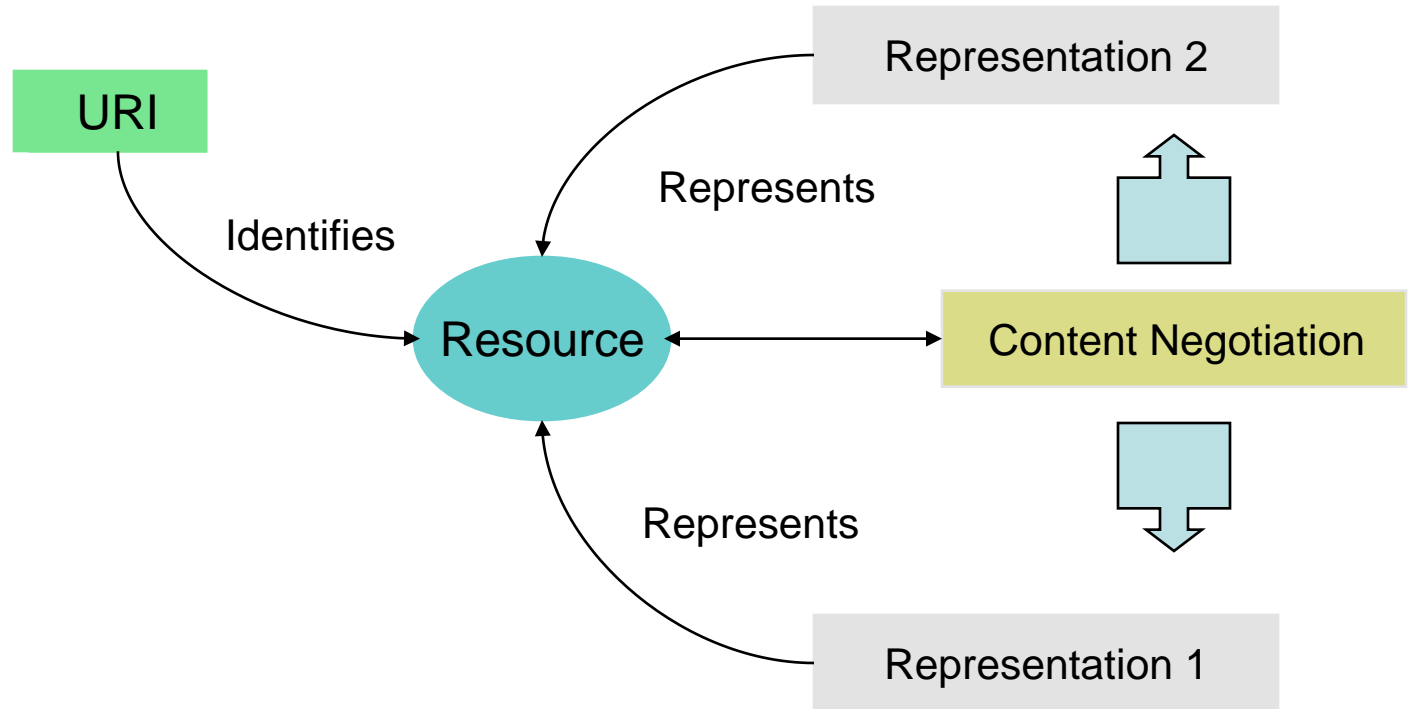
## Core goal of OAI-ORE:

**Facilitate Use and Re-Use of Compound Information Objects (and of their component parts)**

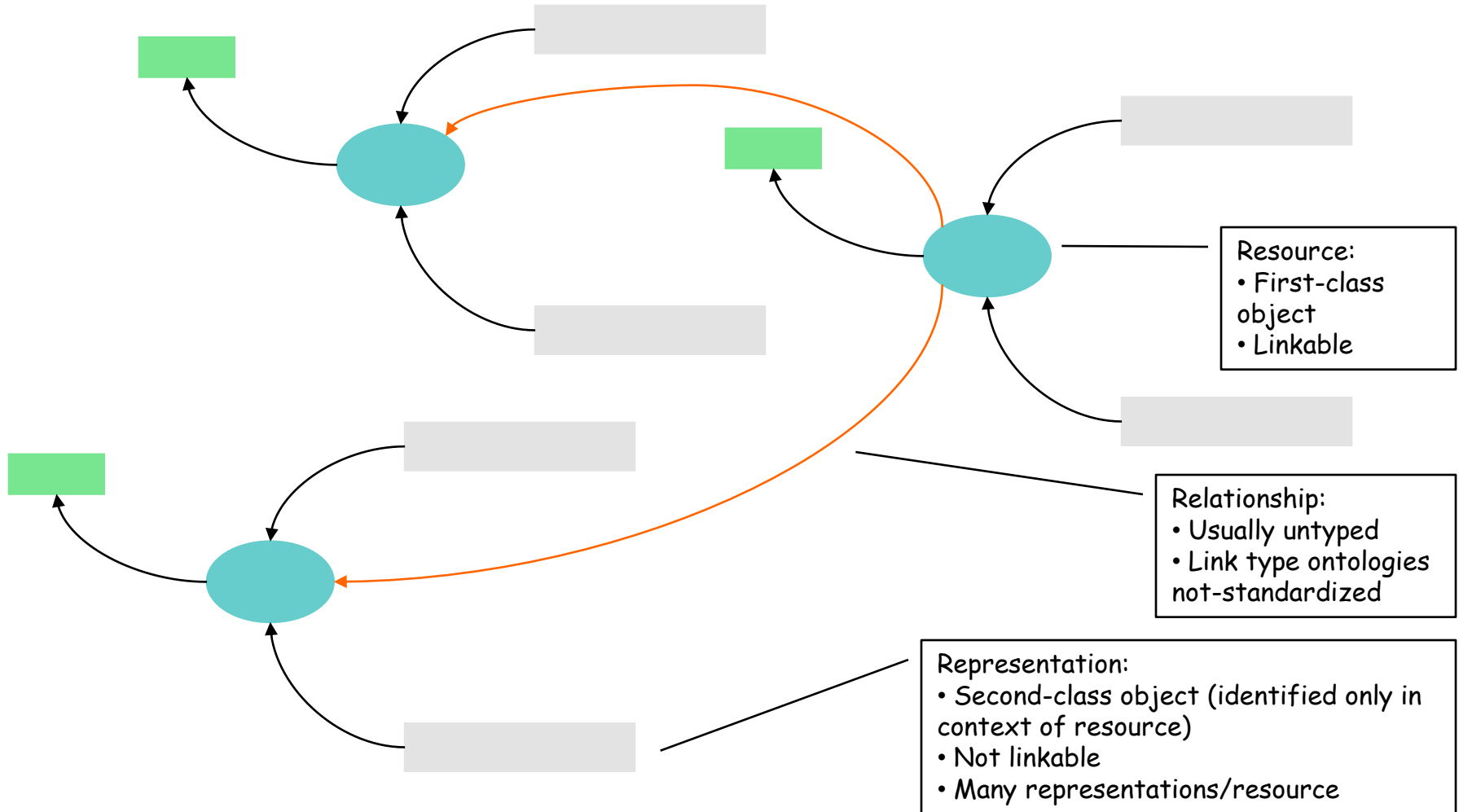
How to deal with compound information objects in a manner that is in sync with the Web architecture?



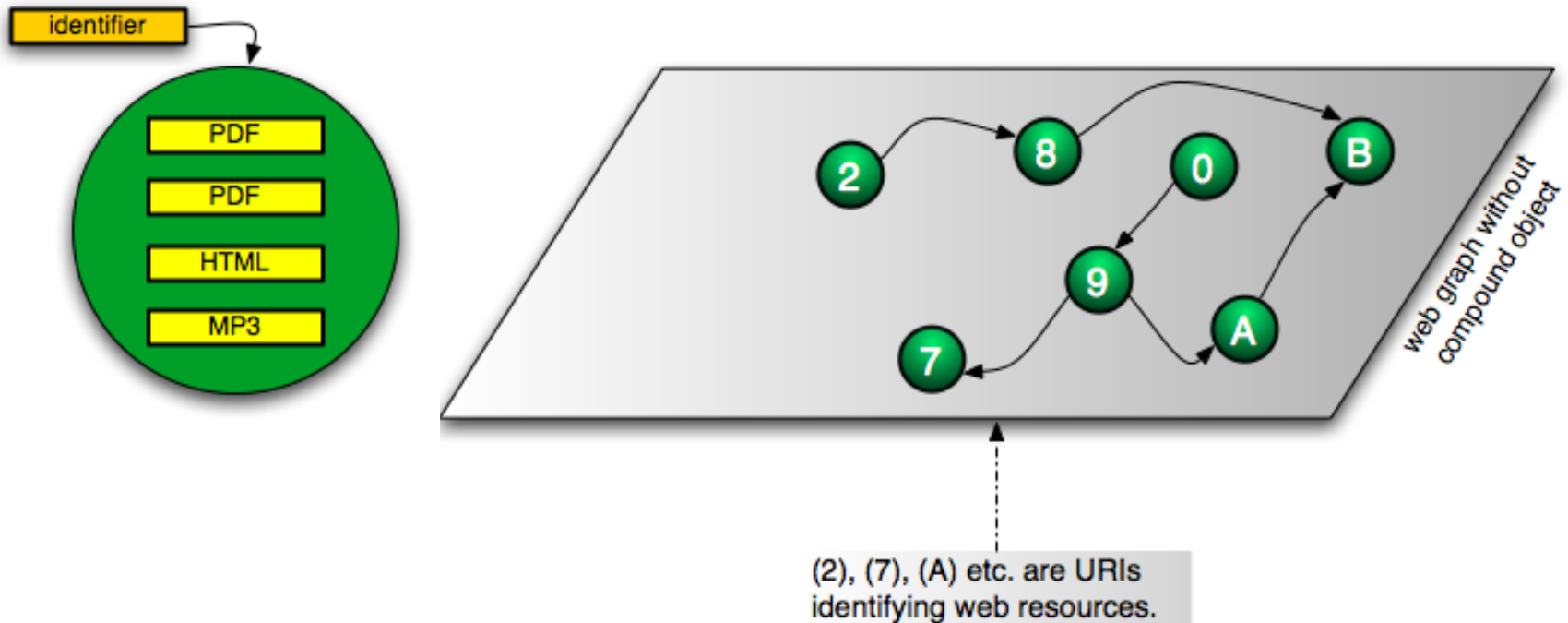
# W3C Web Architecture



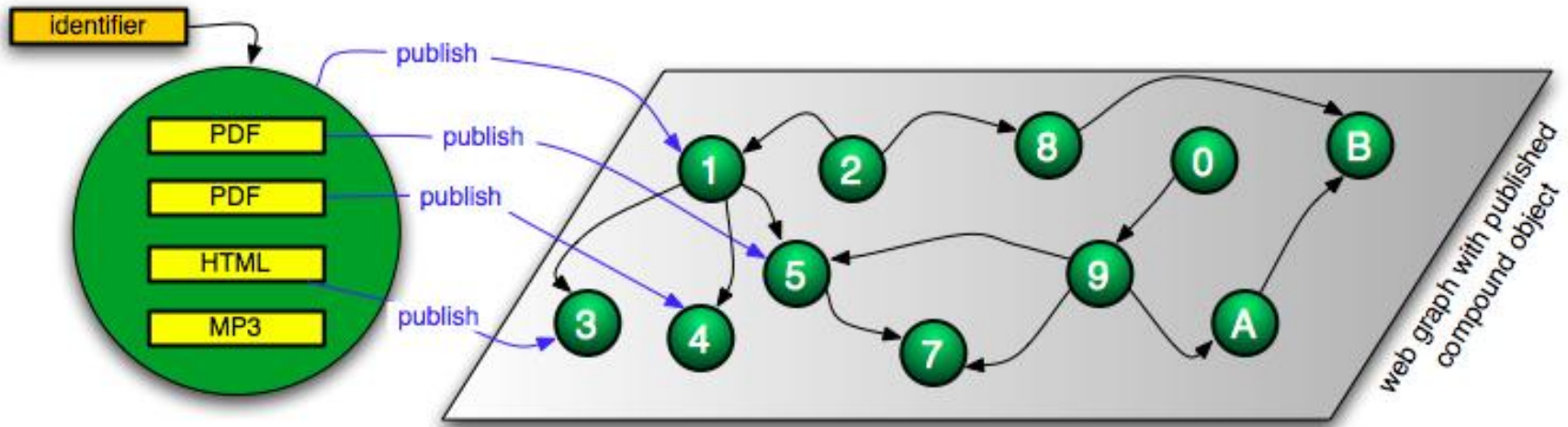
# W3C Web Architecture: more details



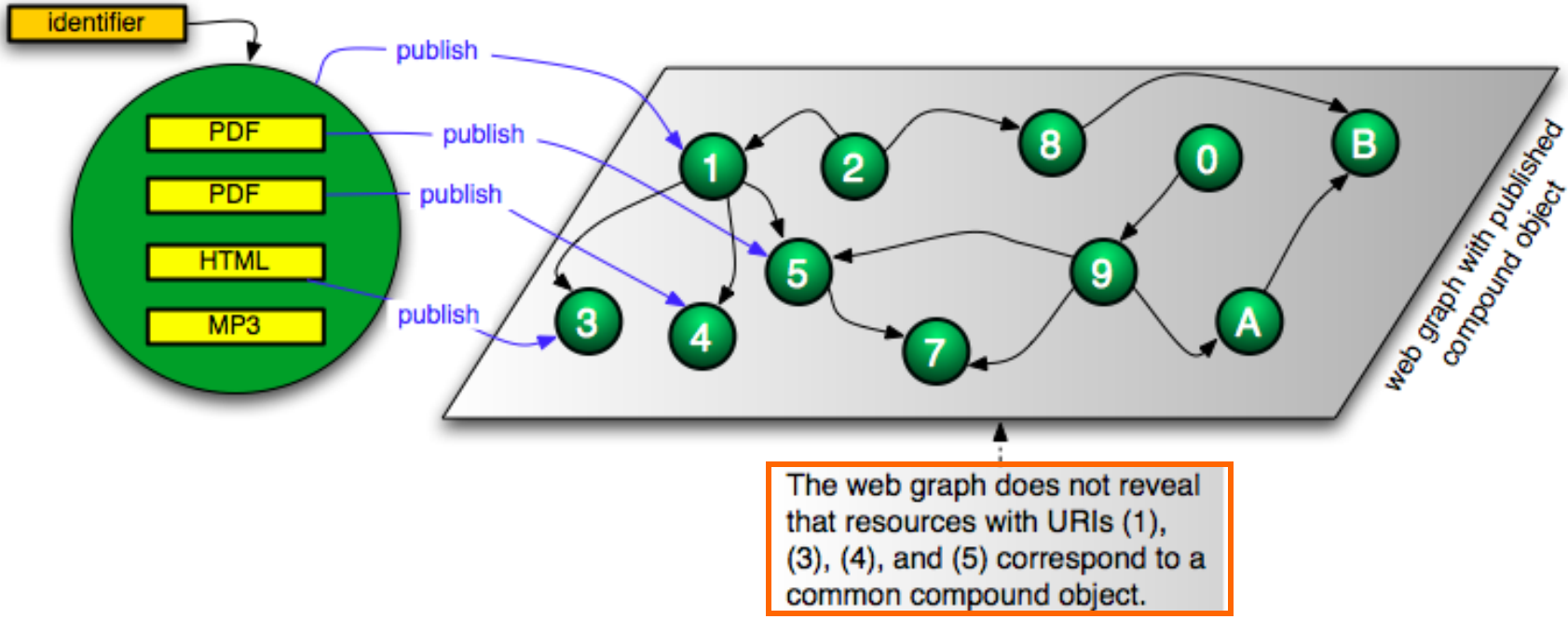
# Publishing a Compound Object to the Web



# Publishing a Compound Object to the Web



# Publishing a Compound Object to the Web: Issues



# Publishing a Compound Object to the Web: Issues

The screenshot shows a web browser displaying the arXiv page for the paper 'Accelerating cosmologies tested by distance measures' (astro-ph/0611775). The page includes a search bar, navigation links, and a sidebar with download options (PostScript, PDF, Other formats) and references. A large orange box highlights a missing image placeholder with the text: 'QuickTime™ and a TIFF (LZW) decompressor are needed to see this picture.'

<http://arxiv.org/ps/astro-ph/0611775>

<http://arxiv.org/pdf/astro-ph/0611775>

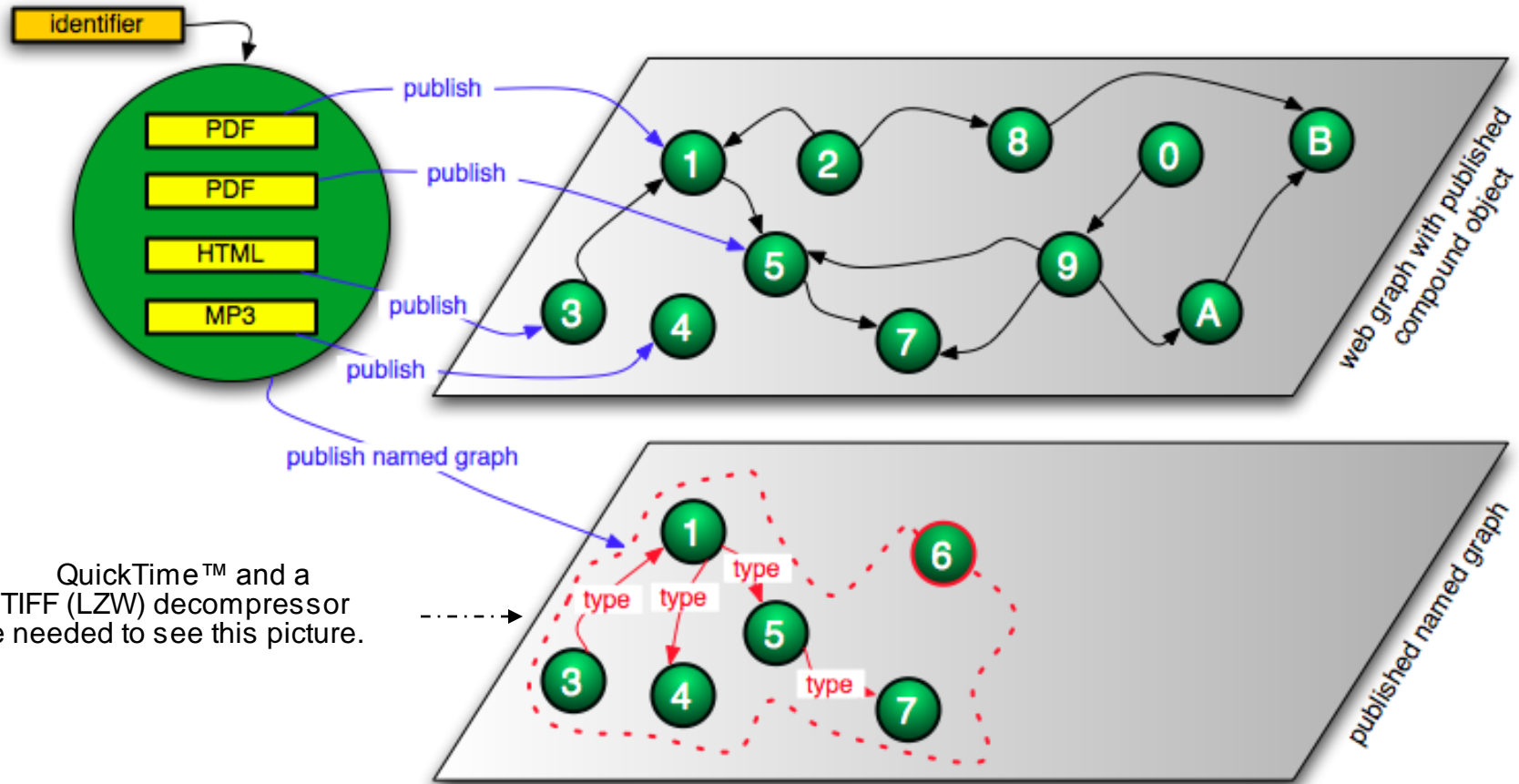
<http://arxiv.org/abs/astro-ph/0611775>

<http://arxiv.org/format/astro-ph/0611775>

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

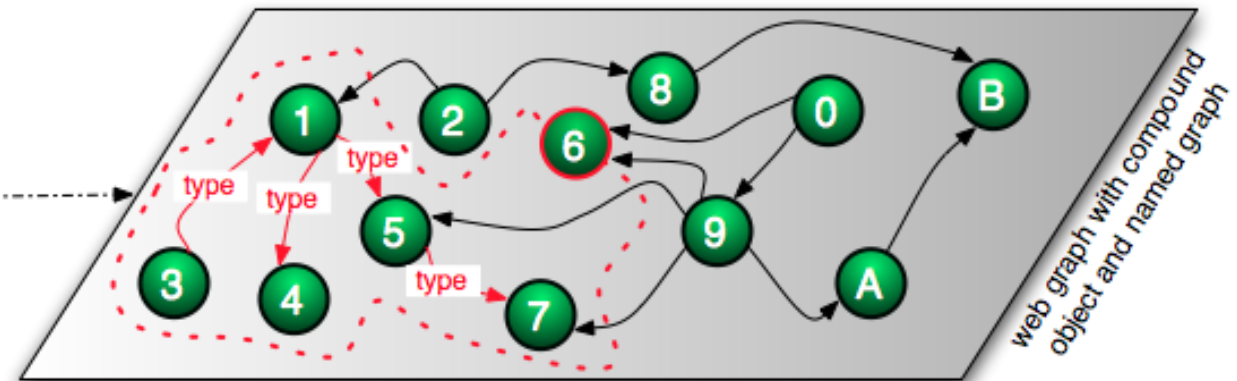


# Publishing a Compound Object to the Web: OAI-ORE

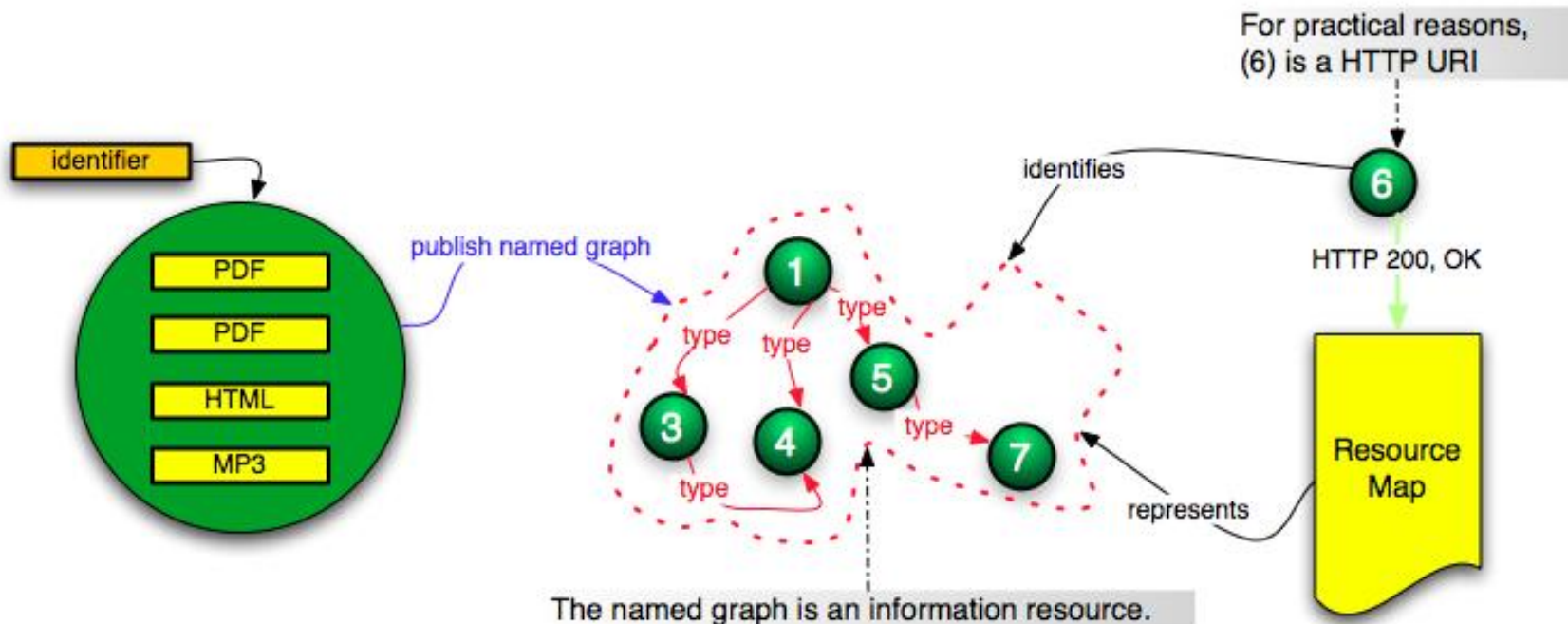


# Publishing a Compound Object to the Web: OAI-ORE

The arc between e.g. the resources with URIs (1) and (5) is typed by means of a URI expressing a relationship type.



# OAI-ORE: Publishing a Named Graph corresponding with a Compound Object



# OAI Object Re-Use and Exchange

## Core goal of OAI-ORE:

**Facilitate Use and Re-Use of Compound Information Objects (and of their component parts)**

**By enriching the web graph with boundary information.**



# OAI Object Re-Use and Exchange

## **Core goal of OAI-ORE:**

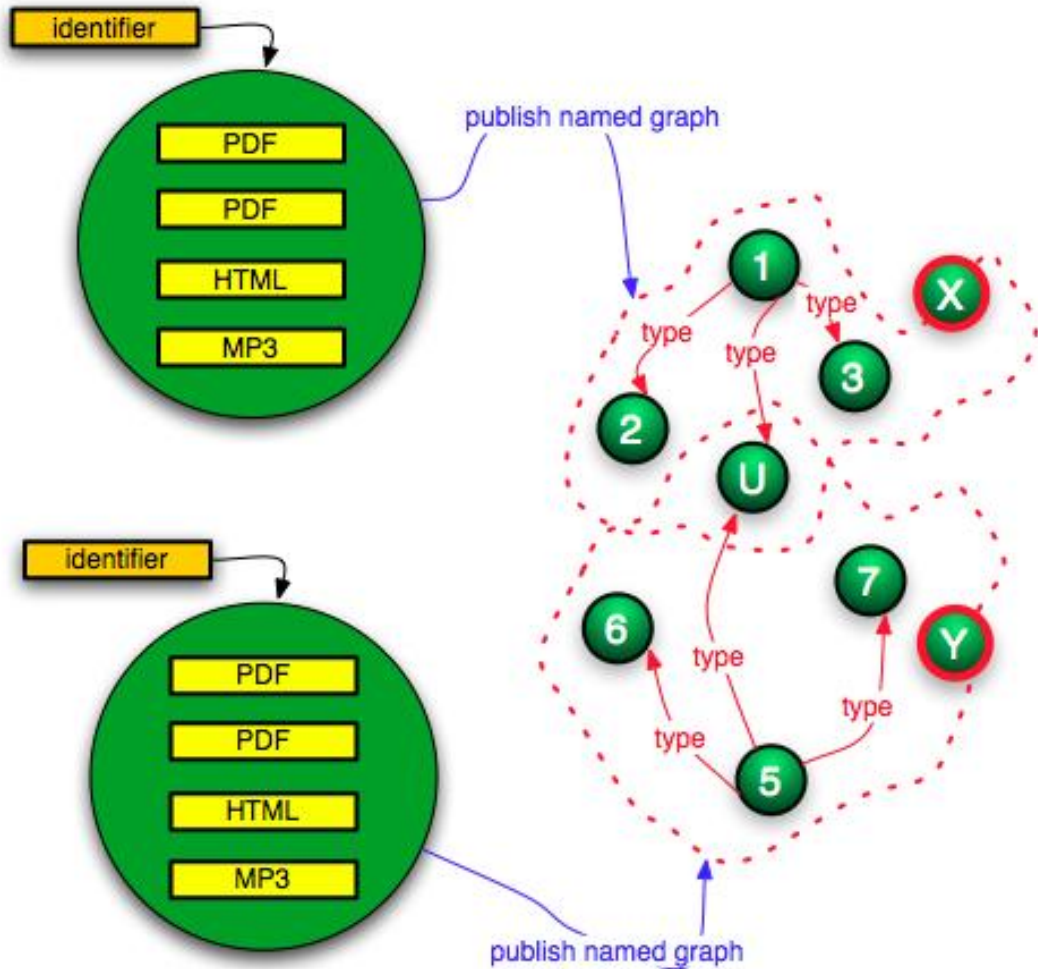
**Facilitate Use and Re-Use of Compound Information Objects (and of their component parts)**

**By enriching the web graph with boundary information.**

**How does this facilitate re-use?**



# Re-Use via URI referencing



Can reference (i.e. re-use) as follows:

(U) : just the resource identified by (U)

(X) : just the named graph identified by (X)

(U) in the context of (X): the resource identified by (U), but as it exists in the context of named graph (X)

# OAI Object Re-Use and Exchange

## **Core goal of OAI-ORE:**

**Facilitate Use and Re-Use of Compound Information Objects (and of their component parts)**

**By enriching the web graph with boundary information.**

What is involved in achieving compound object interoperability using this approach?



# Interoperability layer for compound information objects

The anticipated interoperability layer for compound information objects consists of **approaches to facilitate**:

- a) The **publication of named graphs** to the web as a means to convey compound object (i.e. boundary) information.
- b) **Discovery of these named graphs.**
- c) Assessment of the **trustworthiness of named graphs** as an information source.
- d) Development of a variety of **vocabularies for expressing types of links** between resources denoted by the nodes in a named graph.
- e) Development of a variety of **vocabularies for expressing properties of resources** denoted by nodes in a named graph, especially semantic type, media type, and media format.



# Interoperability layer for compound information objects

The anticipated interoperability layer for compound information objects consists of approaches to facilitate:

- a) The publication of named graphs to the web as a means to convey compound object (i.e. boundary) information.
- b) Discovery of these named graphs.
- c) Assessment of the trustworthiness of named graphs as an information source.
- d) Development of a variety of vocabularies for expressing types of links between resources denoted by nodes in a named graph.
- e) Development of vocabularies for expressing properties of resources denoted by nodes in a named graph, especially semantic type, media type, and media format.

**Bootstrap vocabularies only**



# Regarding (a): Resource Maps

- A Resource Map is the serialization of a named graph that corresponds with a compound object.
  - It is a splash-page for machine consumption
  - Experiments with RDF/XML, TRiX, ATOM, YADS



# Regarding (a): Resource Maps

- Resource Maps **must** allow for simply expressing the resources that are considered part of a compound object.
- Resource Maps **may**
  - Express resources that are not part of a compound object.
  - Distinguish between resources that are part of the compound object and those that are not.
  - Express the relationships among the resources referenced by the named graph.
  - Express the types of the relationships among the resources referenced by the named graph, i.e. label the arcs.
  - Express other information related to the named graph and to the resources that it references such as metadata, etc.



# Regarding (b): Discovery of named graphs

- Class 1: Harvest type discovery
  - Expose Resource Maps via OAI-PMH, RSS, Sitemaps



# Regarding (b): Discovery of named graphs

- Class 2: Linked Data type discovery
  - Convey URI of named graph in HTTP header returned in response to HTTP HEAD/GET against URI of a component of a compound object
    - ORE-specific header:
      - X-OAI-ORE-Named-Graph: <HTTP URI of named graph>
    - LINK header:
      - Link: <HTTP URI of named graph>;  
rel="info:ore/type/named\_graph"
  - Reference a resource in the context of a named graph
    - <a href="http://a.com/foo.html"  
context="http://b.org/foo/rem/">helloworld</a>
    - 



# Next Steps

- Alpha specification by the end of September 2007, covering:
  - Resource Map serialization
  - Discovery
  - Bootstrap vocabularies for relationship types and semantic types
- Test projects building on alpha specifications
  - eChemistry
- Iteration of alpha specifications, ...

