

A non-technical introduction to: **OAI-ORE** (www.openarchives.org/ore/)

Defining Image Access project meeting
“Tools and technologies for semantic
interoperability across scholarly repositories”

UKOLN is supported by:



Julie Allinson <j.allinson@ukoln.ac.uk>
Repositories Research Officer

www.ukoln.ac.uk



Acknowledgements

- This presentation is ***heavily*** based on:
- Report of the Technical Committee Meeting, January 11,12 2007
 - <http://www.openarchives.org/ore/document>
- Presentation at Open Repositories 2007
 - <http://www.openarchives.org/ore/document.pdf>

Overview

- Stands for 'Object Reuse and Exchange'
- Falls within the remit of the Open Archives Initiative, the creators of OAI-PMH
- New piece of work, commenced October 2006
- Funded by the Mellon Foundation, with support from the National Science Foundation in the U.S.
- International focus
- Timescale: 2 years, ends Sep. 2008
- Ideas discussed at the 'Augmenting Interoperability' meeting in New York 2006

Relationship to OAI-PMH

- OAI-ORE is NOT a replacement for OAI-PMH
- OAI-PMH will continue to exist as one approach to interoperability
 - OAI-PMH *metadata-centric*
- OAI-ORE will complement with richer functionality, when this is desirable
 - OAI-ORE is *resource centric*

Relationship to Pathways

- Pathways is about ‘rethinking scholarly communications’
- It proposes an ‘interoperability infrastructure’ comprising a shared data model, surrogate format and 3 shared services
- Pathways provides much of the ‘thinking’ for OAI-ORE and involves the same key players; there is much crossover between the two projects
- OAI-ORE is evolving its own models, concepts and definitions
- Current OAI-ORE work is focussing less on the idea of surrogates and lineage, keys concept in Pathways

ORE project organisation

- Coordinators:
 - Carl Lagoze (Cornell)
 - Herbert Van de Sompel (LANL)
- ORE Advisory Committee
 - UK representatives include Liz Lyon (UKOLN)
- ORE Technical Committee
 - 6 out of 14 members from the UK
- ORE Liaison Group
 - UK representatives include Rachel Heery (UKOLN)

ORE Objectives

- Develop, identify, and profile extensible standards and protocols to allow repositories, agents, and services to interoperate in the context of use and reuse of *compound digital objects* beyond the boundaries of the holding repositories.

Aims (1)

- To provide effective and consistent ways:
 - to facilitate discovery of objects,
 - to reference (link to) objects (and their parts),
 - to obtain a variety of disseminations of objects,
 - to aggregate and disaggregate objects,
 - to enable processing by automated agents

Aims (2)

- To establish the basis for a digital scholarly communication system composed of:
 - systems that manage content such as institutional repositories
 - systems and applications that leverage managed content such as search engines, productivity tools, and data and text analysis services

Compound digital objects (1)

- Content (semantic) types including:
 - Text, *image*, video, audio
 - Datasets
 - Simulations
 - Software
 - Dynamic knowledge representations
 - Machine readable chemical structures
 - Bibliographic and other types of metadata
- Media types including:
 - IANA registered MIME types
 - Other type registries such as the Global Digital Format Registry (GDFR)

Compound digital objects (2)

- Network locations including content from:
 - Institutional repositories
 - Scientific data repositories
 - Social networking sites
 - General web
- Relationships including:
 - Lineage
 - Versions
 - Derivations

Examples

- arXiv paper with different disseminations
- an issue of an overlay journal built from distributed ePrints
- eScience publication combining text, data, simulations
- eHumanities resource combining primary and derived content

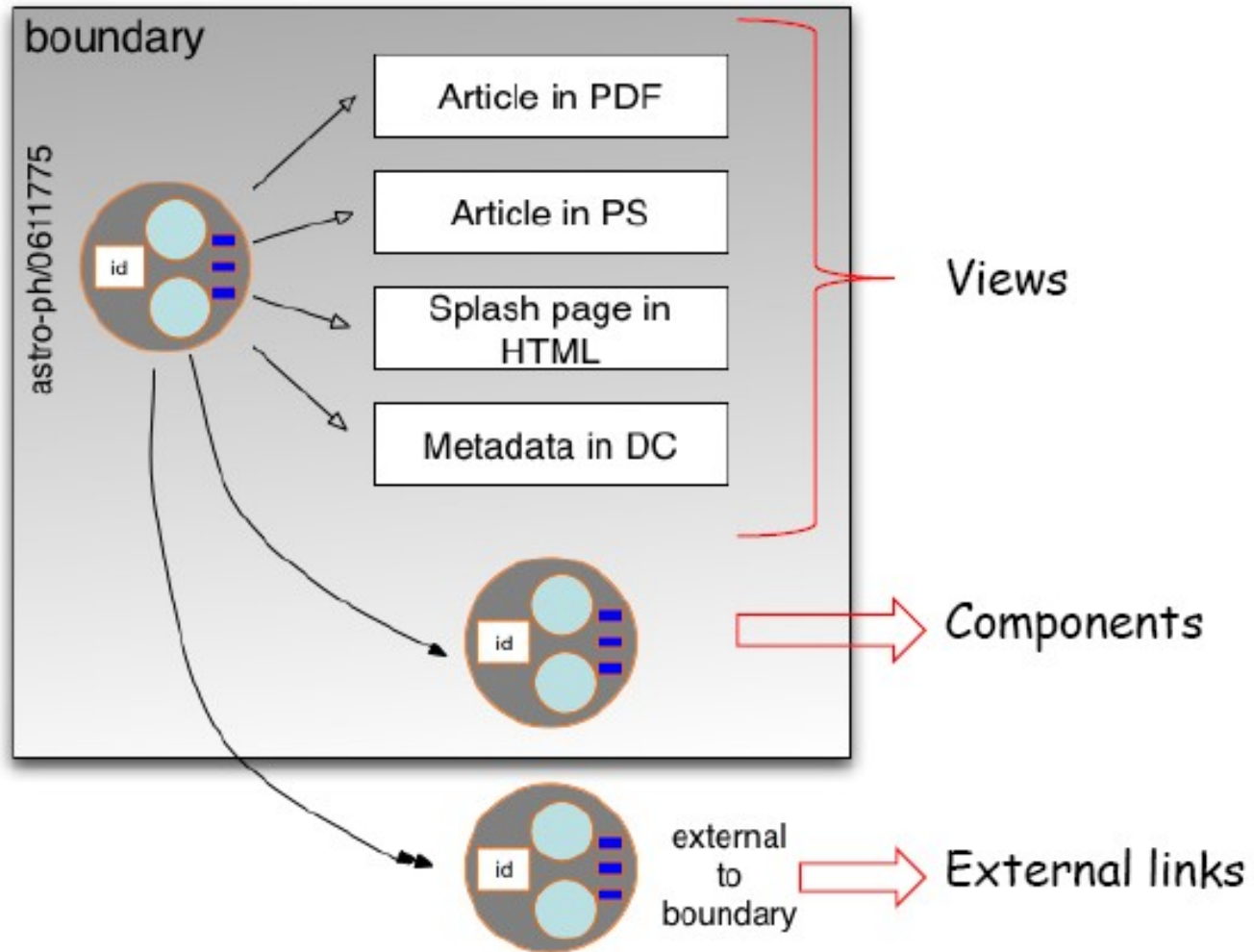
Use cases

- OAI-PRE is working in the context of “workflows supporting research and learning”
- The Technical Committee are developing a set of use cases to illustrate the kinds of workflows and functionality OAI-ORE needs to support

Key concepts

- *Compound digital objects* are “bounded aggregations of *resources* and their *relationships*”
- *Compound digital objects* include:
 - *Views* (alternative ‘presentations’)
 - *Components* (subparts of the main object)
 - both are *Members* of the *Compound digital object*
- It is a requirement that *Compound digital objects*, *views* and *components* can be *unambiguously identified* and *referenced*
- Therefore a *compound digital object*, its *components* and its ‘*views*’ must be modelled as *resources* (as defined by the Web architecture)

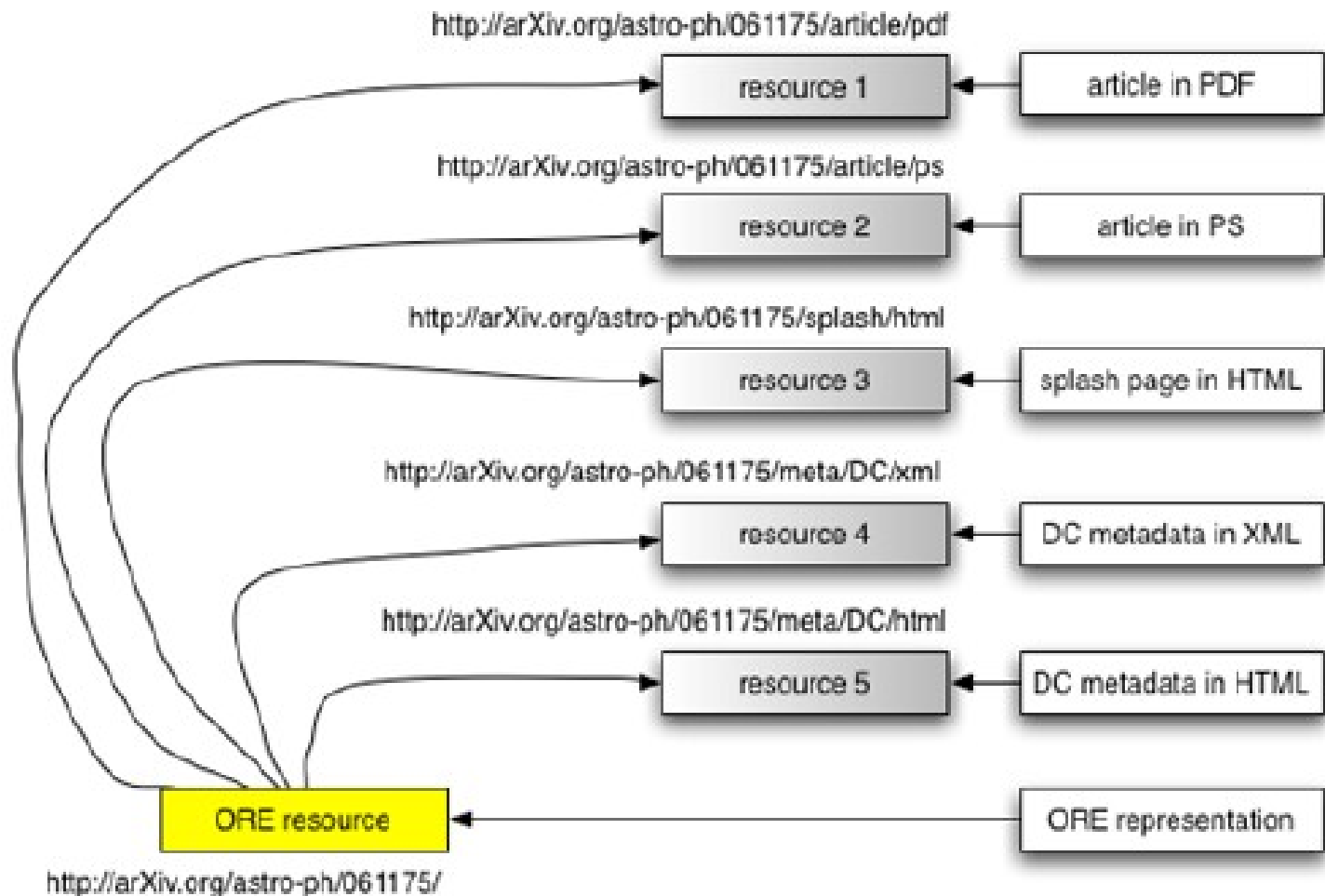
Compound digital object, e.g.



How this is described by ORE

- *ORE resource*
 - “the first-class identifiable object”
 - acts as the access point for service requests upon the aggregation
 - formally expresses the boundaries of the aggregation
- *ORE aggregation*
 - has a defined boundary, as expressed by the *ORE resource*
 - and may have relationships to resources external to the aggregation
 - the *ORE aggregation* is described by the
- *ORE representation*
 - a formal description of the members of the *ORE aggregation* that makes up the *ORE resource*
 - “second-class objects” identifiable only via the ORE resource that they represent

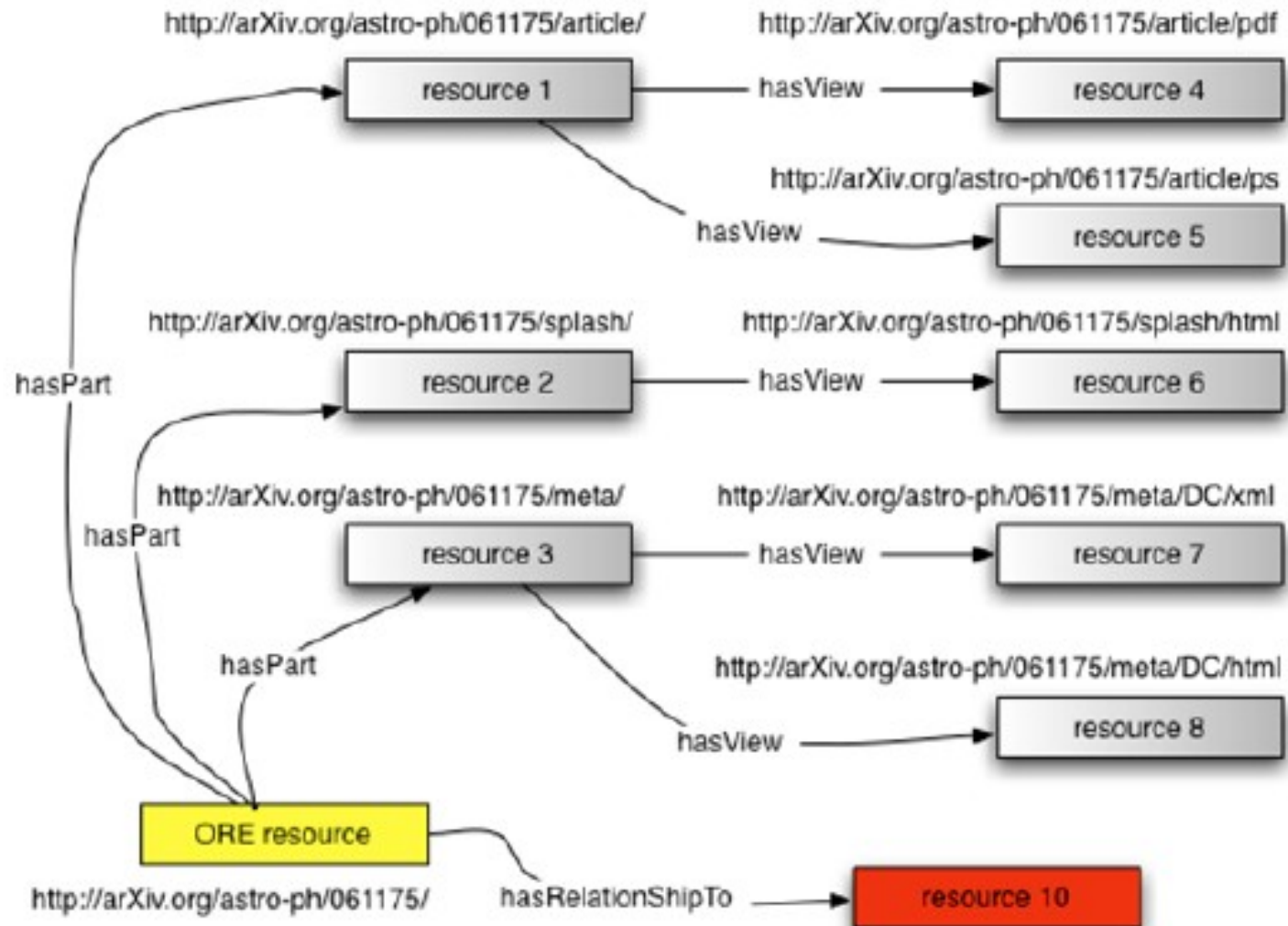
Example modelled according to the ORE model



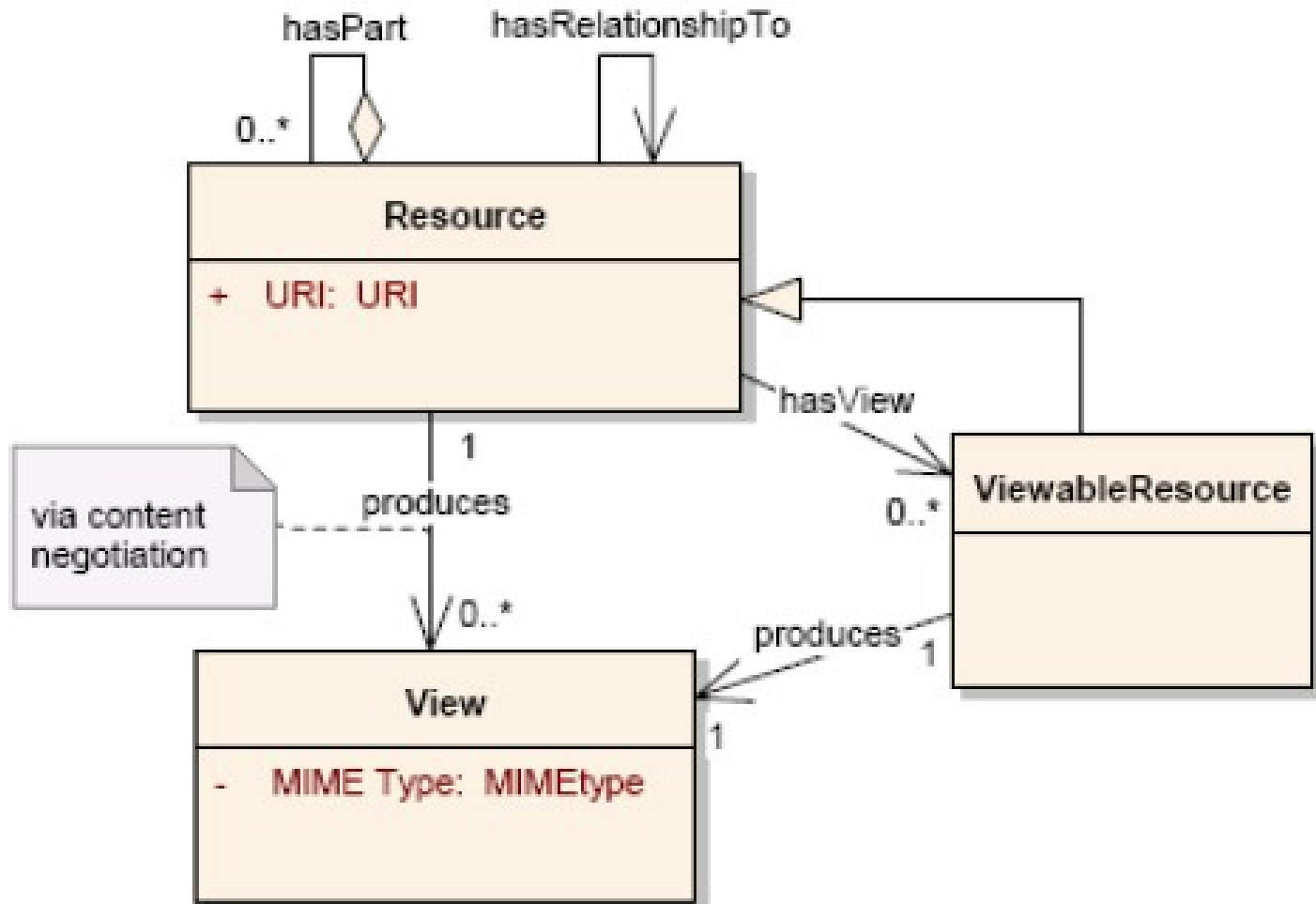
Relationships

- Intra-aggregation relationships
 - *hasPart*
 - *hasView*
- Inter-aggregation relationships
 - *hasRelationshipTo*
 - these may be further defined by community requirements, e.g. lineage, derivation or citation

Example 2, reference to external resource



Preliminary data model



ORE Services

- OAI-ORE Services are transactions that exchange instances of the ORE model
- Three classes:
 - Harvest: a request for a batch of instances that correspond to the ORE model from a set of ORE Resources.
 - Obtain: A request for an instance that corresponds to the ORE Model from a specific ORE Resource.
 - Register: A request to add new nodes or relationships to an ORE aggregation.
- We are not necessarily talking about full asset transfer, *surrogates* may be moved around

Next steps include

- defining the use cases
- reviewing relevant standards for the CaRF (Canonical Representation Format), such as:
 - OAI-PMH
 - RSS/ATOM
 - MPEG 21 DIDL
 - Pathways core

Final disclaimer

- OAI-ORE is very new
- Nothing is finished, finalised or definitive
- The project is being careful about what is disseminated
- But it is likely to have a wide impact