OAI-ORE

a perspective on compound information objects

(www.openarchives.org/ore/)

Defining Image Access final project meeting

Wolfson College, Oxford, 22nd June 2007

Julie Allinson <j.allinson@ukoln.ac.uk>

Repositories Research Officer

UKOLN is supported by:

Supported by

JISC

MLA

Museums, Libraries and Archives Council

www.ukoln.ac.uk

A centre of expertise in digital information management

UNIVERSITY OF BATH

Recap

- Stands for ‘Object Reuse and Exchange’
- Falls within the remit of the Open Archives Initiative, from the creators of OAI-PMH
- Commenced October 2006, 2 year timescale
- Funded by the Mellon Foundation, with support from the National Science Foundation in the U.S.
- International focus, with UK membership on the Advisory Group, Technical Committee and Liaison Group
- Compound Information Objects white paper describes “the web-centric OAI-ORE perspective on compound information objects” … used as a discussion document for May 2007 Technical Committee meeting

UKOLN

Overall 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.

UKOLN

Aims

- 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
- 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

UKOLN

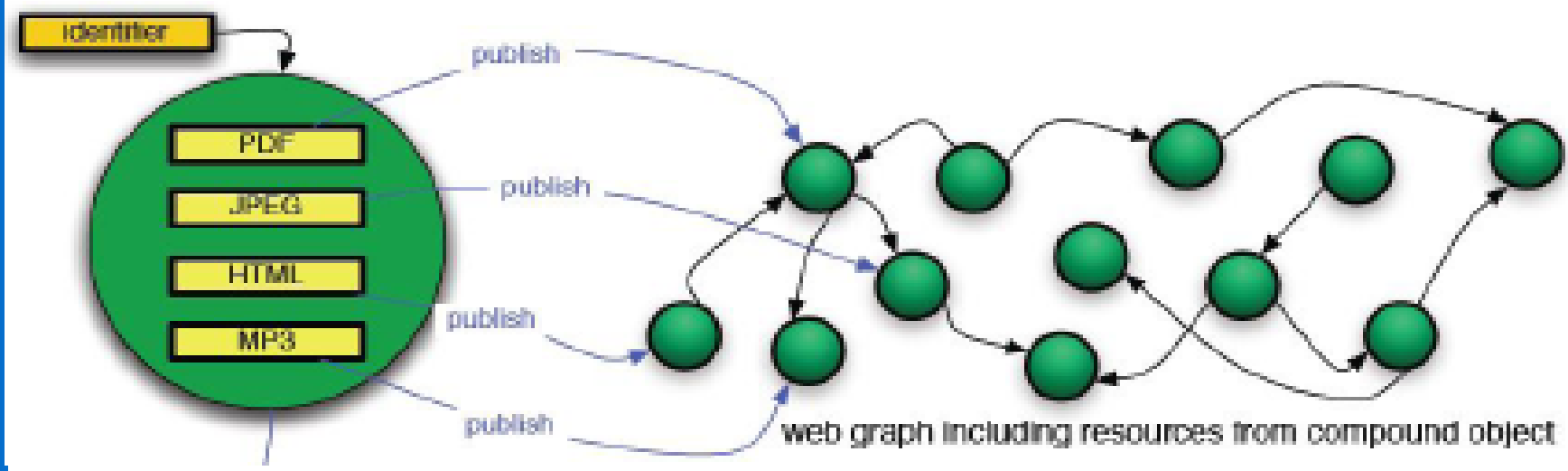
Compound Information Objects

- are aggregations of distinct information units that when combined form a logical whole
- many examples already exist
- information systems can share compound object components using the web architecture
- but the notion of *boundary* and *typed relationships* is lost
- enriching the web with this kind of additional information is the fundamental principle for OAI-ORE

Web architecture

- OAI-ORE is agreed that it must leverage the Web Architecture, through:
  - *URIs* that identify
  - *resources*, which are “items of interest”, that,
  - when accessed through *standard protocols* such as HTTP, return
  - *representations* of current resource state
  - and which are linked via *URI references*

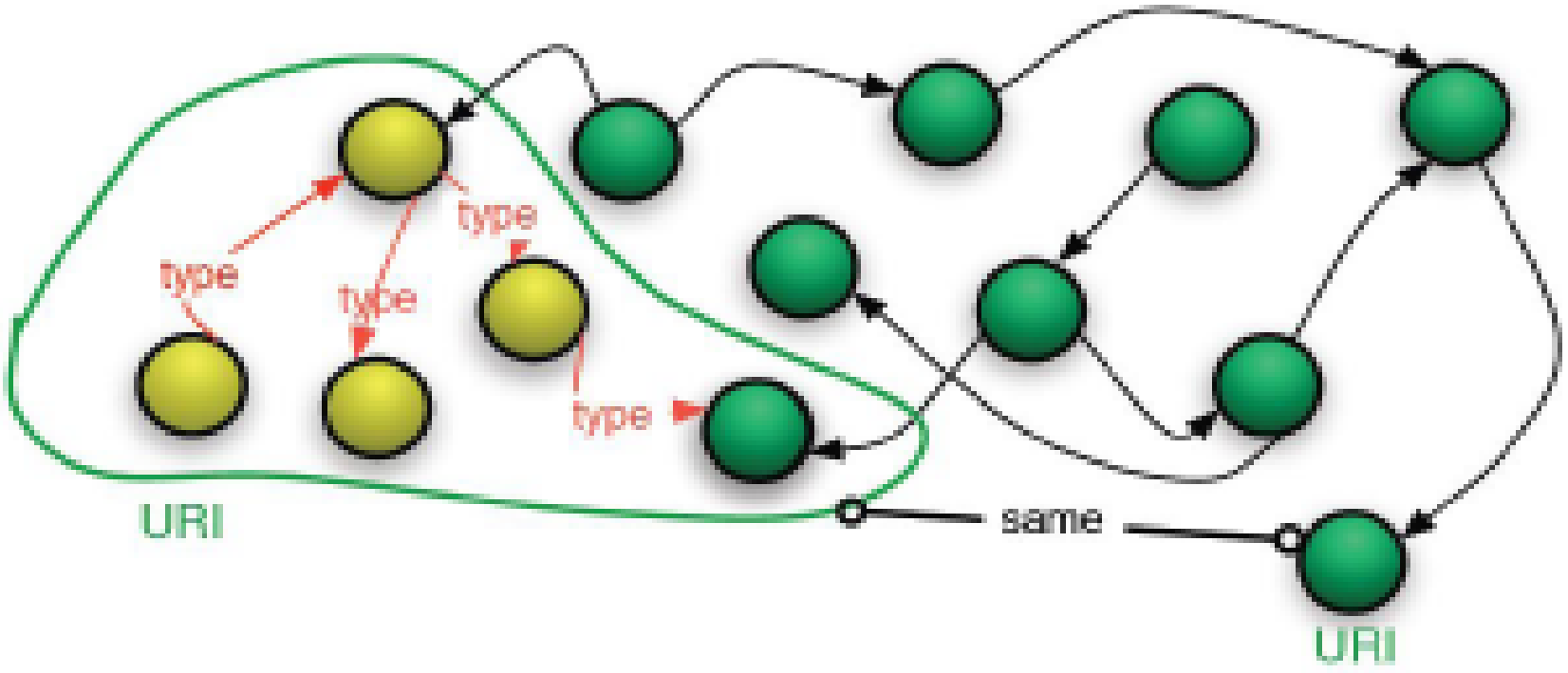
Compound object published to the web



- Compound object and parts have URIs
- Links indicate relationships but cannot show boundaries and true structure in a machine context

UKOLN

Compound object with boundary and typed links

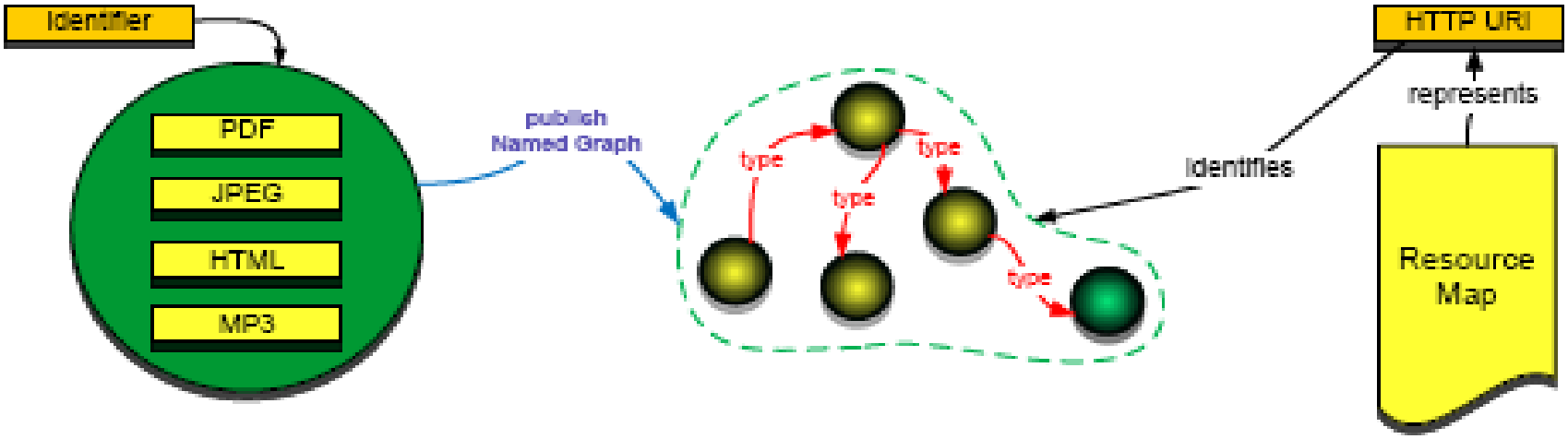


UKOLN

The named graphs approach

- A named graph describes the compound object
- it consists of nodes and arcs where nodes are resources and arcs are typed relationships
- it is a web resource and can be referenced by any resource on the web
- it is not the compound object itself
- it is identified by a HTTP URI
- which points to the Resource Map (ReM) - an encoded description (serialization) of the named graph

Named graph publishing



- Named graph is published at a HTTP URI
- Resource Map is available through content negotiation with that HTTP URI

UKOLN

Outstanding issues

- Rootedness and connectedness – the ‘containment’ node
- Boundaries – internal/external
- Named graph discovery, e.g. through content negotiation
- Authority and ownership, assessment of trustworthiness of Named Graphs
- Development of vocabularies for expressing types of links between resources contained in a named graph
- Development of vocabularies for expressing properties of resources in a named graph, e.g. semantic type, media type and media format

Beyond compound objects …

- Work on a draft specification and serialisation for the ReM is imminent, candidate specifications include
  - RDF/XML
  - TriX
  - ATOM
- OAI-ORE Services are transactions that exchange instances of the ORE model
- These have yet to be scoped out by the project