

Interoperability in the repositories space

Repositories & Preservation
Programme Meeting, 4-5/7/2007

Paul Walk

Technical Manager

p.walk@ukoln.ac.uk

UKOLN is supported by:

JISC

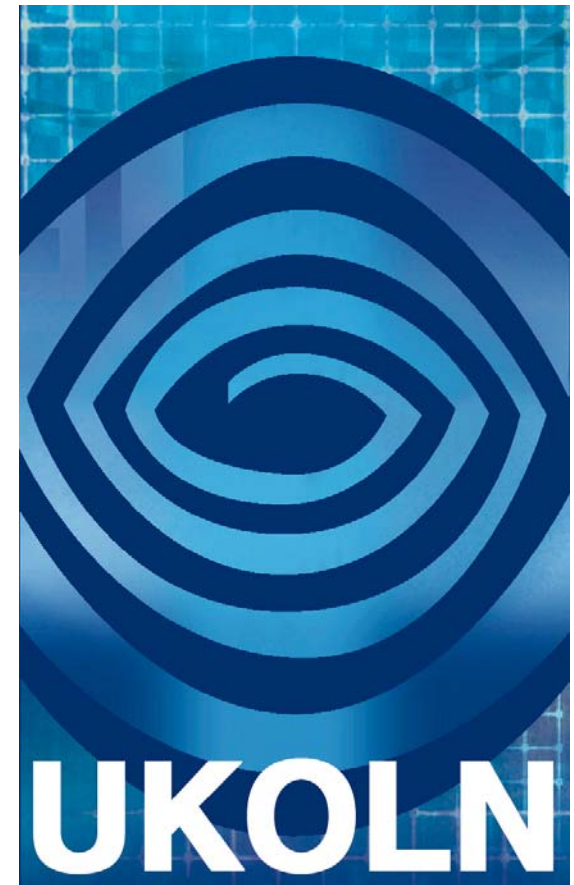
Supported by



Museums, Libraries and
Archives Council

www.ukoln.ac.uk

A centre of expertise in digital information management



repository interoperability (1)

- technical
 - standards & protocols (IE Technical Architecture)
 - interfaces (e-Framework)
- legal/geo-political
 - licensing
 - IPR
 - access management (or not - OAI)
- functional
 - component in a web of repositories
 - component in an enterprise service architecture (SOA)
- semantic
 - thesauri
 - crosswalks

repository interoperability (2)

- repository - repository
 - replication
 - logical repositories, physically distributed
- repository - enterprise system/local application
 - local access control
 - ingest - 'smart deposit'
- repository - shared infrastructure services
 - name authority
 - thesauri
 - federated access management
- repository - remote web service
 - cross search
 - annotation/tagging

JISC Information Environment

- “The aim of the Information Environment is to help provide convenient access to resources for research and learning through the use of resource **discovery** and resource **management** tools and the development of better services and practice.”

JISC Strategy 2007-2009

- in a Web 2.0 world a hybrid approach is probably needed:
 - discovery needs to accommodate external, commercial resources
 - the **management** aspect needs to be integrated

inside and outside the institutional boundary

- repositories on the web
 - all about access - APIs and simple interfaces
 - networked - data and social interaction
- repositories in the enterprise
 - information/data/content is a **core business asset**
 - it needs to be **managed**

the Web 2.0 effect on access (& contribution!)

- important, but probably exaggerated in importance
- has demonstrated what can be achieved with a few fundamental, open standards:
 - HTTP
 - XML
 - RSS
 - HTML
- what can we learn from this?
 - a philosophy of 'just good enough to get the job done' can work in certain circumstances
 - this approach works best in simpler environments
 - might not fit the enterprise/institution

the mighty Atom?

- Atom has two formats:

1. the syndication format

- alternative to RSS, but very similar in function
- technically better, more strictly defined (e.g. dates)
- attempt to escape woes of a dysfunctional open standards community

2. AtomPub

- previously abbreviated as APP
- conceived as standard format and protocol for publishing to blogs & wikis typically
- wide interest in adoption for publishing to CMS systems in the general sense
- recent selection for the SWORD project
- increasing adoption despite not yet having been ratified (usually a good sign!)

engagement with the enterprise

- the repository is now recognised as a separate functional component (service)
- SOA (the e-Framework)
- the line between repositories and content management systems will blur:
 - e.g. *Alfresco* - CMS with many repository ideas built-in
 - JSRs for content management being implemented commonly in these areas
 - VREs as enterprise-grade software for the research institution?
- institutions might want to start to treat their learning materials as a core business asset
 - MIT took the opposite view - but who has followed them?
- separate the *management* of the stuff from the *use* of the stuff

community resources & developments

- JISC IE Technical Architecture
 - <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/>
- JISC RRT & RSP
- Open Archives Initiative
 - <http://www.openarchives.org>
- Common Repository Interfaces Working Group (CRIG)
 - identifying common scenarios & use-cases
- e-Framework
 - from the scenarios we identify in real use, we should be able to extrapolate service usage models (SUMs)
- Object Reuse and Exchange (ORE)
 - <http://www.openarchives.org/ore/>
- Knowledge Exchange
 - Netherlands, Germany, Denmark & UK
- + don't forget the JISC Programme!

JISC Standards Catalogue

- now ready to use
 - <http://standards.jisc.ac.uk>
- 6-monthly update cycle - snapshots
- snapshots can be cited in bids and circulars
- part of a three-fold contextual model
 - policy
 - **summary of standard (including risk assessment)**
 - compliance (enforcement??!)
- issues which need to be addressed:
 - feedback - how?
 - long editorial cycle (6 months) - how are minor revisions dealt with?
 - reconciling authoritative, time-stamped versions with current reality

enough interoperability already....

...time to show some *interoperation!*