An introduction to …

OAIS and reference models for repositories

JISC Repositories and Preservation Programme Meeting
24-25 October 2006, London

Julie Allinson

Repositories Research Officer

UKOLN, University of Bath

UKOLN is supported by:

Supported by

Museums, Libraries and Archives Council

www.ukoln.ac.uk

www.bath.ac.uk

A centre of expertise in digital information management

Overview

- Reference models, context and background

- OAIS, purpose, use

- OAIS for repositories

"Reference models, why bother?"

- Bill Olivier

What *is* a reference model? (1)

- A framework for understanding significant relationships among the entities of some environment, and for the development of consistent standards or specifications supporting that environment. A reference model is based on a small number of unifying concepts and may be used as a basis for education and explaining standards to a non-specialist. (OAIS 2002, p. 1-12)

What *is* a reference model? (2)

- An abstract task model of what has to be accomplished to meet the needs addressed, described in a way that is independent of how it is accomplished
- The description of the chosen means of implementing this model, including:
  - The roles and activities that humans and computer systems are respectively to play in accomplishing this task.
  - The workflow or processes involved.
  - The use cases involved at one or more points in the process that form part of the computer system requirements at each point.
  - From the use cases are derived:
    - a specification of the service or services called on, together with links to the specifications and bindings used
    - a specification of how the various services are co-ordinated in those cases where they have to work together.

Oliver, Roberts and Blinco 2005, p. 8

www.ukoln.ac.uk

UKOLN A centre of expertise in digital information management

Levels of abstraction

- High level [informative]
  - Domain maps, models, ontologies, communication tool
  - OAIS is a high-level reference model
- Medium level
  - Processes, practices, workflows, scenarios, use cases, context and systems
- Low level [normative]
  - Specifications, orchestration / choreographic models, reference implementations, architectures
  - NB: The E-Framework now talks of Service Usage Models rather than Reference Models

Reference models : what for?

- Reference models are communication tools, they offer a shared vocabulary and a shared basis for understanding
    - between domains
    - developers, repository implementers
- Offer repositories a tool for gap analysis
- DO guide practice, DON'T mandate practice
- … should evolve to reflect practice
- But …
    - There are existing software platforms for repositories … don't want to retrofit a reference model to what we already know/do
    - Is there any value in focusing purely on the abstract?
    - This may distract us from real problems that need to be solved.

UKOLN A centre of expertise in digital information management

Existing reference models, architectures and frameworks

- OAIS <www.ukoln.ac.uk/repositories/digirep/index/OAIS>
- JISC IE Discovery to Delivery (D2D) Reference Model <www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/dlf/>
  - draft for discussion, by Andy Powell, based on the …
- DLF Service Framework <www.diglib.org/architectures/serviceframe/>
- JISC reference models projects <www.elearning.ac.uk/del/refmodels>
  - COVARM, FREMA, eP4LL, LADIE, SCRI
- DELOS Digital Library Manifesto – in development
- CORDRA (Content Object Repository Discovery and Registration/Resolution Architecture)
- aDORE architecture
- More … see Dan Rehak's presentation at the JISC-CETIS Conference <www.e-framework.org/events/conference/programme>

OAIS : an overview without *the* diagram

- Reference model for an Open Archive Information System
- Developed by the Consultative Committee for Space Data Systems (CCSDS)
- Provides a framework for the standardisation of **_Long-Term Preservation_**
- OAIS environment contains
  - Producer, Consumer, Management and the OAIS archive/repository
  - Designated Community is a key concept
- The OAIS contains
  - Information Model (SIPs, AIPs, DIPs)
  - Functional Model (Ingest, Archival Storage, Data Management, Administration, Preservation Planning, Access)
  - Minimal set of mandatory responsibilities
  - Additional sections on preservation techniques and interoperability

OAIS : observations

- OAIS is not prescriptive, compliance requires adherence to a minimal set of requirements
- Uses neutral (or fairly neutral) terminology
- Facilitates communication across different domains (e-science, archives, digital libraries, institutional repositories etc.)
- Further work on certification and trusted repositories is being done by RLG (in the US) and the DCC (in the UK)

Useful for repositories

- OAIS already exists (no re-invention)
- It is used and is becoming more widely used, it has a community
- It is actively developed, in an open forum
  - ISO standard; 5 yearly reviews
- It is relatively easy to comply with OAIS
  - asks only that preservation is considered and planned for
  - allows 'long-term' to be defined by the repository
- Offers a high-level, abstract guide on the functions a repository should carry out and the structure of the information it handles
  - Not a benchmark as such, but a tool for analysis and comparison of systems and practices
- Facilitates communication

But …

- It is focussed on Long-Term Preservation, perhaps NOT the primary business requirement for a repository

- It is a long document and can seem overly complex (forthcoming revisions might help)

- As a conceptual model, it doesn't help with implementation

- It doesn't go far enough in reflecting the increasingly complex ecology and the interactions between repositories and between repositories and external services

- Attempts to 'certify' may be a barrier to compliance for smaller repositories

Is OAIS relevant for JISC projects?

- OAIS is an approach, not the **only** approach
- OAIS can provide a high-level model for repositories
    - For communication it helps to ensure we are 'on the same page'
- Repositories must consider their full range of business requirements – OAIS may not help here
- OAIS may be used in conjunction with other levels of abstraction
- JISC projects and programmes are helping to document these, e.g.
    - e-Framework Service Usage Models
    - Scenarios, use cases, workflows documented by projects
    - Reference model projects

Final thoughts …

- For *all* repositories preservation and access can and should co-exist … with caveats
  - Repositories must consider preservation but may opt not to do it
  - They may get a specialist preservation service to do it for them
  - They can define their own 'long-term'
  - (but not everybody agrees with this view)
- OAIS could have a very positive role in heightening preservation awareness
- OAIS-lite documentation might help smaller repositories become preservation-aware
- Forthcoming work on the Repository Ecology will examine OAIS in relation to interoperability issues
- … and its relationship with the Information Environment and the e-Framework

UKOLN A centre of expertise in digital information management