ARROW Update

Presentation to CAUL
15 September 2005

Geoff Payne,
ARROW Project Manager

The ARROW Project is funded by the Australian Commonwealth Department of Education, Science and Training, under the Research Information Infrastructure Framework for Australian Higher Education.

arrow.edu.au

The ARROW Consortium comprises Monash University [lead institution], National Library of Australia, The University of New South Wales and Swinburne University of Technology.

arrow
australian research repositories online to the world

MONASH University | NATIONAL LIBRARY OF AUSTRALIA | UNSW | SWINBURNE UNIVERSITY OF TECHNOLOGY

ARROW - Summary of design criteria

- A generalised institutional repository solution for research information management
- Initial focus on managing and exposing traditional “print equivalent” research outputs
- Expanded to managing other digital research outputs
- Design decisions accommodate management of other digital objects such as learning objects and research inputs such as large data sets
- DEST Research reporting and audit, and Research Quality Framework likely to drive deposit of content by academics and research managers in ARROW universities

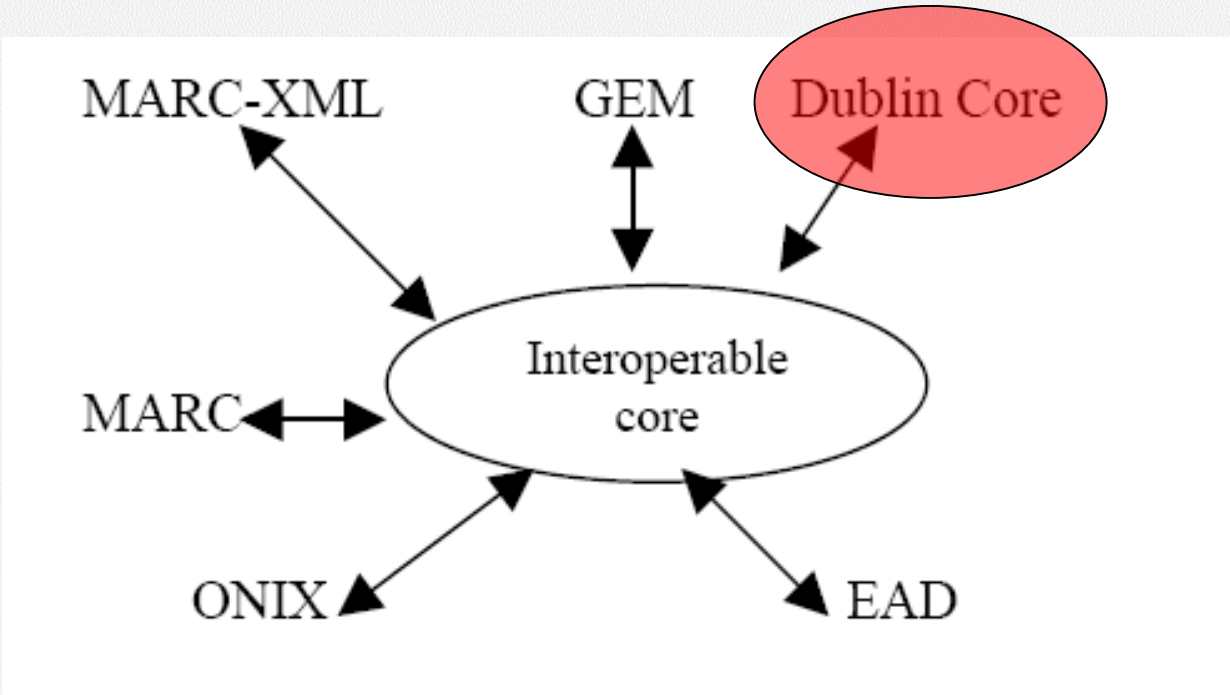
ARROW media types

- Known:
  - Digital objects ARROW can store and render (e.g. display, run, manipulate) For example
    - PDF, Raw text, still images, moving images
    - Associate Disseminators for custom behaviour rendered by the server
- Unknown
  - ARROW will store the bits and allow them to be retrieved, but will rely on the user to provide any software required to render the content, for example
    - MSOffice documents
    - Data sets

ARROW Metadata Strategy

- Supports metadata schemata to suit individual data models
  - No requirement to shoehorn all metadata into one schema
  - Each stored object can retain metadata developed for it by the community of practice which generated the object
  - Maintains flexibility to store many types of digital objects in the repository
  - No need to anticipate every object type now
  - Maps metadata to Dublin core to populate the ARROW Discovery Service

OCLC Metadata Interoperability Core



From: Godby, Smith and Childress. 2003. “Two paths to interoperable metadata” p. 3 at http://www.oclc.org/research/publications/archive/2003/godby-dc2003.pdf

ARROW stages

- Demonstration (2004)
  - Developing architecture, selecting, testing and developing software
- Deployment (late 2004 – end 2005)
  - Populating the ARROW Partners' repositories
- Distribution (mid 2005 – end 2006)
  - Enabling others to participate
    - Under review for earlier participation by others

ARROW: an Information Management Tool

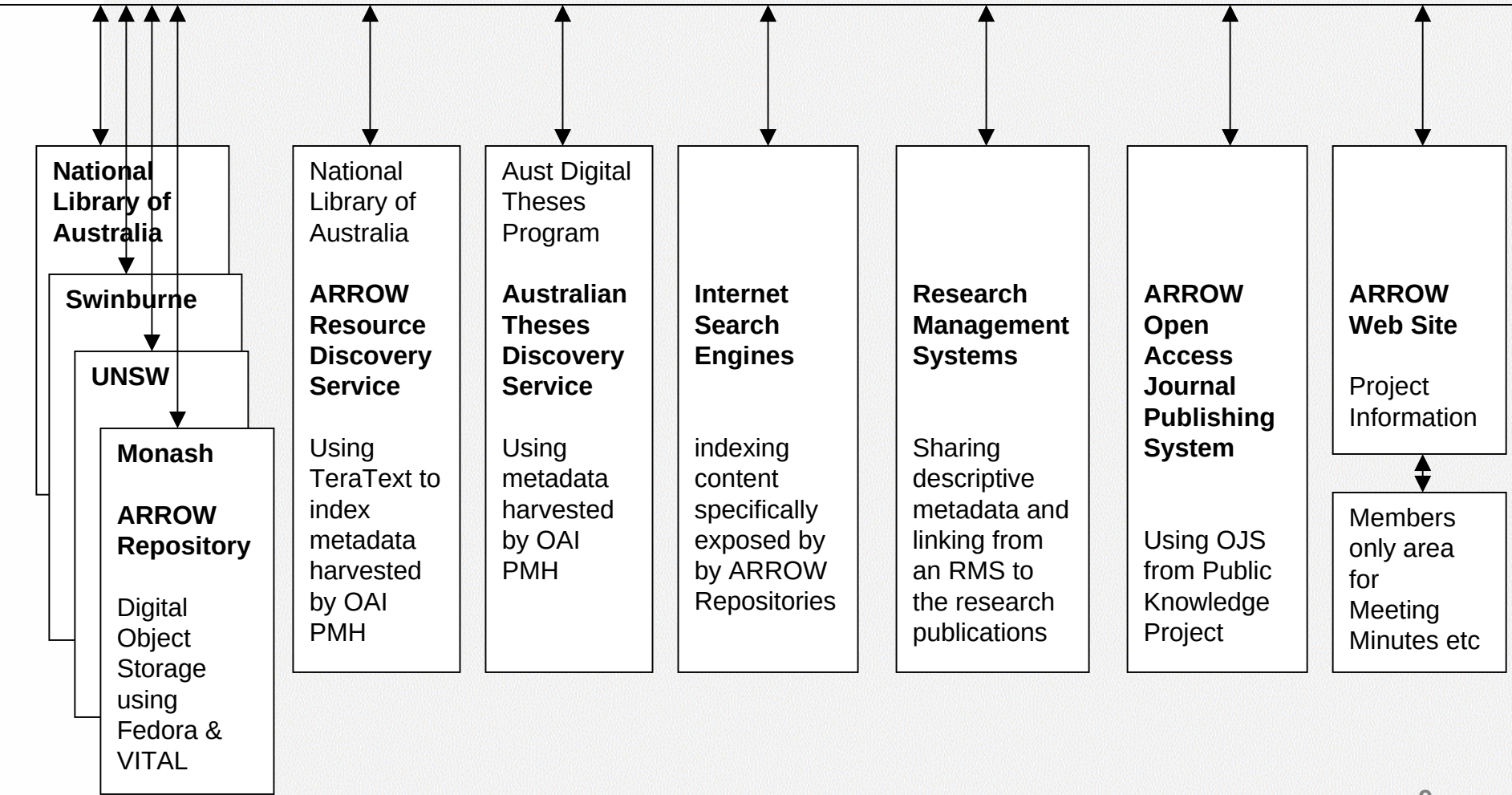
DEST Research evidence

- Around 30-40% of Australian university research funding comes from government
- At present an annual statistical return is required and audit evidence of research outputs is compiled as collections on paper
- ARROW can improve the efficiency of this process as an information management tool
- In the Monash context this will capture 4000 publications annually

Research Quality Framework

- Australian Government wanting
  - wider exposure of Australian research
  - Demonstrable value for money, “research impact”
- Moving to a research quality framework similar to the United Kingdom Research Assessment Exercise (RAE)
- RAE 2008 guidelines include assembling research outputs in institutional repositories
- RAE utilises approx 70 discipline specific expert panels to rank research outputs from UK universities
  - Around 25% of UK govt research funding goes to the top 4 Universities
- ARROW is envisaged as a tool for the Australian RQF

ARROW Branded Services Profile



ARROW progress to date – Open Journal Publishing

- Provides full suite of publishing functionality
  - Peer review management
  - Assembling and publishing journal issues
  - Well liked by academics using the software
- Swinburne University is leading ARROW open journal publishing activities
  - E-Journal of Applied Psychology launched 1 July 2005
  - See http://www.swin.edu.au/lib/ir/onlinejournals/ejap/
- Informed by University of Technology Sydney “Portal” e-journal
  - See http://epress.lib.uts.edu.au/journals/portal/

ARROW progress to date – Resource Discovery Service

- Harvesting metadata from a variety of repositories
  - Eprints.org
  - DSpace
- ARROW repositories not yet being routinely harvested
  - Awaiting Fedora support for OAI Sets (Due in Fedora 2.1)
    - To allow selective harvesting to
      - Resource Discovery Service
      - Picture Australia
      - Australian Digital Theses Program
- http://search.arrow.edu.au

ARROW progress to date - Repositories

- "Paper content models" (actually a large spreadsheet!) pending software management of content models
  - To ensure consistent metadata and datastreams for like objects
- VITAL 1.3 software release on production repository servers
- VITAL 2.0 alpha over Fedora 2.0 in testing at ARROW and VTLS
- Awaiting Fedora 2.1 software release due 30 September to incorporate XACML access control into VITAL by January 2006
- Work plan for July 2005 to May 2006 agreed between partners and with VTLS
  - Enhanced user interface first priority in VITAL 2.1 due October

ARROW & DEST Research Reporting and Audit

- Design work proceeding on an interface between Research Master (RM) and ARROW for gathering DEST research evidence
  - Monash, Swinburne, UNSW all use RM v.4, but the solution will be generalised to accommodate other practices

ARROW Software Development – Current Status August 2005

- VITAL 2.0 alpha over Fedora 2.0
  - Image Management
  - Text Documents in PDF
  - VITAL Batch ingest tool for digital objects and/or their metadata
  - Handles integration for automatic assignment of persistent identifiers
  - SRU/SRW interface
  - Audio, Moving Pictures support
  - Support for Google spidering

VITAL 2.0 capabilities

- Manual or batch ingest of digital objects by partner staff
  - Automatic assignment of Handles persistent identifiers
  - JHOVE content validation
  - MARCXML metadata to Dublin Core transformation
  - Advanced searching
  - SRU/SRW
  - OAI harvesting to populate the ARROW Discovery Service
  - User configured indexing
  - User defined Fedora object structures
  - Web submission tool for end user e-theses deposit

VITAL 2.0 capabilities

- Web based deposit for theses
  - Followed by review and manual ingest by staff
  - By November, cloning this for images, journal articles, books, book chapters, working papers, conference papers pending software development for generic content model management
- Batch ingest tool
  - Match metadata files and object files on various criteria
- Exposure of content to web search engines
- VITAL 2.1 will be based on Fedora 2.0
  - Improved user interface including browsing
- VITAL 3 to include integration with Fedora 2.1
  - Support for XACML access controls
  - Support for OAI Sets for metadata harvesting

ARROW demo – VITAL 2.0 alpha

http://arrowdev.lib.monash.edu.au:8000/access

http://arrowdev.lib.monash.edu.au:8000/access/explorer.php

http://arrowdev.lib.monash.edu.au:8000/cgi-bin/etd/submit.cgi

http://arrowdev.lib.monash.edu.au:8000/cgi-bin/etd/review.cgi

http://arrowprod.lib.monash.edu.au:8000/access

ARROW demo – VITAL 2.0 alpha

Monash theses ingested using the web self submit tool:

http://arrowdev.lib.monash.edu.au/hdl/1959.100/630 Treloar, A.E : Hypermedia online publishing: the transformation of the scholarly journal

http://arrowdev.lib.monash.edu.au/hdl/1959.100/628 9.7 MB Robilliard, Frederick Emile:Studies of hollow-cathode metal vapour ion lasers

Monash theses ingested using the batch ingest tool:

http://arrowdev.lib.monash.edu.au/hdl/1959.100/1104 Smith, David Alan, Antecedents and outcomes of multiple dimensions of accountants' organisational commitment

http://arrowdev.lib.monash.edu.au/hdl/1959.100/1108 Wilson, Campbell:Visual information retrieval via inference networks

http://arrowdev.lib.monash.edu.au/hdl/1959.100/1111 Karmakar, Gour Chandra: An integrated fuzzy rule-based image segmentation framework

ARROW demo – VITAL 2.0 alpha

JPEG Images

http://hdl.handle.net/1959.100/459 Composite weather image of a tropical cyclone Creator NASA

http://hdl.handle.net/1959.100/457 Satellite image of Victoria and Northern Tasmania Creator NASA

MrSid images with navigation

Advanced Search:Title: ag050002 http://hdl.handle.net/1959.100/516 Victoria Dock, 1972 and 2002

Advanced Search:Title: ag050009 http://hdl.handle.net/1959.100/507 Victoria Dock, circa 1910 and 1942

ARROW demo – VITAL 2.0 alpha

Text and supporting images

Advanced Search:

Title: muddies http://hdl.handle.net/1959.100/418 History Australia, Volume2, No.1, 2004. Ferals and their muddies: Making a home in the bush

Text only

RTF text Advanced Search:

Title: requirements http://hdl.handle.net/1959.100/486

ARROW demo – VITAL 2.0 alpha

XML plus images: Advanced Search: Title: residential
http://hdl.handle.net/1959.100/283 Melbourne 2030: Chapter 5 - Residential infill and its threat to Melbourne's liveability.

MPEG movie

http://hdl.handle.net/1959.100/586 Medical computer animation #20

Quiktime movie

http://hdl.handle.net/1959.100/566 Medical computer animation #10

mp3 audio

http://hdl.handle.net/1959.100/571 Ash Grunwald, Bakelite Radio and Blues Progression 5Mb

AVI movie: Advanced Search: Title: fantastic

http://hdl.handle.net/1959.100/551 Fantastic Four, movie trailer **WARNING large file 16Mb** DivX codec must be installed first to view.

ARROW demo – VITAL 2.0 alpha

ARROW Branding of Access Portal

http://arrowdev.lib.monash.edu.au:8000/access/?view=ARROW

Building on ARROW

August 2005 Strategic Infrastructure Initiative funding announced for (among others):

- DART (Monash University as lead institution)
  - Supporting the e-research lifecyle
  - Includes managing large datasets in the ARROW repositories
    - Interfacing Fedora and Storage Resource Broker or similar technologies
    - Managing annotations
- RUBRIC (University of Southern Queensland as lead institution)
  - Evaluating ARROW as part of identifying repository solutions for regional universities in Australia and New Zealand
    - Application to management of learning objects
- IP management in repositories (Queensland University of Technology lead institution)
  - Including Creative Commons Australianisation

Summary

Functionality In Hand:

- VITAL Manager can ingest content and metadata edited externally with XMLSpy – Not for the casual user
- Web submission for theses
- Batch Ingest matching metadata and digital objects
- Access portal for searching
- Access Explorer for specifying indexing

Still to come:

•Imminent

- Web ingest for other content types
- Enhanced user interface with browse capabilities
- RM4 interface

•Early 2006

- XACML Access control at Object and datastream levels
- Support for OAI Sets for metadata harvesting

•Mid 2006

- Generalised content model management

Questions?

Further information?

Details of the ARROW project can be found at:

arrow.edu.au

The ARROW site includes links to the FRODO projects and a glossary of repository acronyms and projects