

Ingest: selected issues

Initiatives to link articles and datasets

Project StORe (Source To Output Repositories)

JISC-funded, multi-partner project investigating the ability to move seamlessly from source (data) repositories to output (publications) repositories. The project completed mid-2008

Repository for the Laboratory (R4L) Project

<http://r4l.eprints.org/>
JISC-funded project applying repository technology to experimental data capture, analysis and reporting processes in chemistry to enable linking between datasets and articles and between related datasets. Project partners: University of Southampton and ALPSP (Association of Learned and Professional Society Publishers)

Dublin Core

<http://dublincore.org/>
Managed by the Dublin Core Metadata Initiative. The Dublin Core metadata set consists of 15 standard elements that are repeatable and freely orderable. Qualified Dublin Core includes additional levels of detail to the basic set

MARC (MACHine Readable Cataloguing)

<http://www.loc.gov/marc/>
Library of Congress-maintained standard for bibliographic metadata. Some regional variants exist, though these are tending to become standardised into USMARC, the US version

METS (Metadata Encoding and Transmission Standard)

<http://www.loc.gov/standards/mets/>
The METS schema is a standard for encoding descriptive, administrative, and structural metadata regarding objects within a digital library, expressed using the XML schema language of the World Wide Web Consortium. The standard is maintained in the Network Development and MARC Standards Office of the Library of Congress, and is being developed as an initiative of the Digital Library Federation

MODS (Metadata Object Description Schema)

<http://www.loc.gov/standards/mods/>
Metadata Object Description Schema (MODS) is a schema for a bibliographic element set that may be used for a variety of purposes, and particularly for library applications. The standard is maintained by the Library of Congress

SWAP (Scholarly Works Application Profile)

<http://www.jisc.ac.uk/whatwedo/programmes/reppres/swap.aspx>
JISC-funded project to develop an application profile for Dublin Core. Article on SWAP: <http://www.ariadne.ac.uk/issue50/allinson-et-al/>

Complex object metadata schema

DIDL (Digital Item Declaration Language)

Developed by the Moving Picture Experts Group to accurately describe multimedia objects. ISO Standard
Used by LANL

IMS Global Learning Consortium Content Packaging XML Binding

http://www.imsglobal.org/content/packaging/cpv1p1p2/ims_cp_bindv1p1p2.html

SCORM (Shareable Content Object Reference Model)

<http://www.adlnet.gov/>

CCSDS (Consultative Committee for Space Data Systems) XML Formatted Data Unit (XFDU) Structure and Construction Rules

<http://public.ccsds.org/publications/MOIMS.aspx>
Published September 2008

Deposit interface

SWORD (Simple Web Service Offering Repository Deposit)

<http://www.swordapp.org/>
Lightweight protocol for depositing content from one location to another. A profile of the Atom Publishing Protocol (ATOMPUB). Enables batch deposit
JISC-funded project managed by UKOLN

Repository-to-repository exchange (work to date is mostly, but not exclusively, for preservation purposes)

'Mining for ORE': Repository Developers' Challenge at Open Repositories 08 (supported by the JISC CRIG [Common Repository Interface group])

The developer team from the PRESERV2 Project won this competition with a deposition of an entire Fedora repository (metadata plus content) into EPrints and vice versa, implementing the OAI-ORE standard
Beta release: <http://www.openarchives.org/ore/0.9/primer.html> Video at <http://blip.tv/file/866653>

SHERPA DP2

<http://www.sherpadp.org.uk/index.html>
JISC-funded project, extended to a second phase, that investigated the provision of preservation services available for institutional repositories. The Arts & Humanities data Service developed a high-level model and shared preservation environment for use by these institutions. The second phase extended the study to cover repositories that archive additional, wide-ranging content types and implement different organisational models. Project partners SHERPA and the AHDS

TIPR (Towards Interoperable Preservation Repositories)

http://blogs.fcla.edu/index.php/digitalarchive/2008/09/11/news_from_the_fda_september_11_2008
Two year (beginning October 2008) project funded by the Institute of Museum and Library Services, carried out at Florida Center for Library Automation at Cornell University and New York University Libraries. The project will test the exchange of information between three repositories (based on DAITSS, aDORe and DSpace). It will use the ECHO DEP profile and build upon MathArc work

CRIS-repository interoperability

CRIS are (generally relational) databases recording the context of research within an organisation. Recently, CRIS have moved towards web-based system, based on defined data models. The main standard is CERIF (Current European Research Information Format), developed by EuroCRIS and recommended by the EU. There are some exploratory initiatives currently exploring the issues involved in creating interoperability between CRIS and Open Access (literature) repositories and an increasing dialogue between the CRIS and OA communities on this

AIHT (Archive Ingest and Handling test)

A project of the NDIIPP. An experiment to provide better understanding of digital archiving experiences at institutional level. Drew many institution-pertinent conclusions and two general ones: (i) preservation is an outcome and a variety of measures to preserve are better than one (ii) data-centric approach is preferable to tool-centric or process-centric. D-Lib article on the project: <http://www.dlib.org/dlib/december05/shirky/12shirky.html>

MathArc: Ensuring Access to Mathematics Over Time

<http://www.library.cornell.edu/dlii/MathArc/web/index.html>
Projects partners Cornell University and Goettingen University. Created a system for preserving and disseminating mathematics and statistics literature by interchanging information packages between geographically and administratively separated databases using OAI-PMH

PRESTA (PREMIS Requirements Statement)

<http://www.apsr.edu.au/presta/index.htm>
Australian project undertaken by APSR. Studied how PREMIS could be used to transfer content from one repository to another.

ECHO DEpositary

<http://www.ndiipp.uiuc.edu/>
NDIIPP project at the University of Illinois Urbana-Champaign, 2004-9. Implemented instances of four different repository software applications and designed an interoperability architecture. The architecture consists of a METS interoperability profile and a 'hub' service that translates to and fro from the profile