DepositMOre: Applying tools to increase full-text content in institutional repositories

Steve Hitchcock, Tim Brody, JISC DepositMOre Project
ECS, WAIS, University of Southampton

#rspevent, Increasing the full-text deposits in your institutional repository, London, 12 June 2013
In this talk ...

• What’s mine is yours, what’s yours is mine - sharing ‘full-text’ content between repositories and services?
• DepositMO user tools and SWORDv2 (recap)
• DepositMOre repository tools
• DepositMOre repositories
• Roles for repositories. Is MOre content the primary target?
DepositMO

Impenetrable acronym – changing the *Modus Operandi* of repository deposit June 2010 – Jan 2012

DepositMOre

Self-explanatory goal – getting more stuff into your repository Sept 2012 – Mar June 2013

MOre

MO
Credits and acknowledgements

• PI: Les Carr
• Development: Tim Brody, Dave Tarrant
• MOre repositories and ‘depositors’
  – University for the Creative Arts: Anne Spalding
  – Goldsmiths: James Bulley
  – Leeds: John Salter
  – Glasgow: William Nixon
  – Southampton (archaeology): Gareth Beale, Hembo Pagi, Graeme Earl
• Me
Why SWORD?

Simple Web service Offering Repository Deposit

- SWORDv2 (2012) items can be created, updated, replaced, or deleted (CRUD).
- SWORDv2 endpoints built into current versions of EPrints (3.3) and DSpace (3.1)
SWORD deposit interface

Facebook SWORD client
Watch Folder – a file manager-based deposit interface
From the Desktop to the Cloud: Leveraging Hybrid Storage Architectures in your Repository

David Tanant, Tim Brody, and Leslie Carr

Add-­in deposit client

http://www.eprints.org/depositmo/
DepositMO/re tools

**Watch Folder**: extending the user interaction model for intelligent bulk ingest to deal with 100s of documents, involving original user testers from within arts and archaeology, communities that generate large volumes of non textual digital objects.

**Word add-in**: impact of a content creation tool on deposit rates will be less immediate.

**EasyChair Deposit Tool**, lists a user’s authored items in Easychair, which hosts over 15,000 conferences. The tool checks if these items are present in the user’s selected repository. Any that have not been deposited can be added with one click.

Enhancing and testing repository deposit interfaces, OR12, [http://eprints.soton.ac.uk/340997/](http://eprints.soton.ac.uk/340997/)
Choice of content targets

- EasyChair – conference papers
- Microsoft Academic Search
- Europe PubMed Central (RJ Broker – EDINA)
- arXiv
- YouTube/Vimeo – video
- Archaeology image generators (PTM)
MOre repositories: chosen tools

UCA
Goldsmiths
Glasgow

Leeds
Southampton (archaeology)
Southampton University on YouTube

University of Southampton Economics graduate
by sotoncomms • 9 months ago • 438 views
Alexandra Raikes, graduate in BSc Economics talks about the support she has had during her studies and her future plans.

University of Southampton - Southampton engineers a Raspberry Pi Supercomputer
by sotoncomms • 7 months ago • 368,496 views
Computational Engineers at the University of Southampton have built a supercomputer from 64 Raspberry Pi computers and Lego...

Broderick Fine Foods tasting at the University of Southampton
by Broderick FineFoods • 1 month ago • 437 views
We invited students on campus at the University of Southampton to try our most popular product, Lovington's Ice Cream. What did...
Importing YouTube video to EPrints
Importing YouTube video: paste URL

Paste data

http://www.youtube.com/watch?v=cdKoUjhyBl
Importing YouTube video: item type auto-identified
Importing YouTube video: simple metadata auto-added
Importing YouTube video: done
Vimeo video deposited in GRO

http://eprints.gold.ac.uk/8284/
Feedback from UCA on video tool

• One good example Andy Joule Annimilus
  http://www.research.ucreative.ac.uk/1270/1/Annimilus.mov
• showed our VC and two deans UCA Research Online record which included an imported video using the import tool for vimeo and they were very happy at the way in which the video played immediately from the repository.
• There is however at least 15-20 further items that could be added
• The benefits are likely to be ease of use of importing the films, and increased visibility and discoverability.
• works in Firefox but not on IE. I understand that the issue might well be resolved on EPrints 3.3. we are on v3.2.
• it would be useful to use the tool to upload additional information to existing records.

Anne Spalding, Mar-Apr 2013
Polynomial Texture Mapping (PTM)/Reflectance Transformation Imaging (RTI)

- Reflectance Transformation Imaging (RTI)  

- Processing output formats
  - Camera-raw -> digital negatives -> JPEG (for processing) -> RTI (using RTI Builder) -> PTM archive

- Archaeological applications of polynomial texture mapping: analysis, conservation and representation  
  [http://eprints.soton.ac.uk/156253/](http://eprints.soton.ac.uk/156253/)
Unzipping a PTM archive: unaided

EPrints 3.3 test server – no PTM app
Unzipping a PTM archive: with a research data app and the PTM app

EPrints 3.3 test server – with ReCollect data deposit app and PTM app
Spotting PTM unzip tool
Unzipped: 3 files (it could be 80!)
Users’ views on EPrints PTM app

• The system is almost there. There are some usability issues, but I guess most of them are related with the eprints itself, not the ptm app. The final unzipped content page needs a bit of work, as at the moment you cannot access original files one by one, only by downloading whole zip.
• I think the main benefit of this system will be when actual XML file can be used to get all the metadata and have all the form fields prefilled. Sadly those XMLs do not exist yet but can be automatically generated, which we will. At the moment there are too many steps from upload to deposit.
• We've started to collect data tho, and it will be a massive task (I think we have almost 1000 entries) which I think requires bulk upload functionality.

Hembo Pagi, Southampton archaeologist

• It would be ideal if the ingestion process stripped out the thumbnail and showed that on the archive page and then elsewhere on the page showed the snapshots, perhaps in a gallery view.

Graeme Earl, Southampton archaeologist
New Digital Library at Leeds

Based on the EPrints platform ...

will be producing a large number of digital objects, mainly contained in folders that represent the object being digitised, e.g. several TIFFs representing individual pages of a manuscript would be held in one directory. For this 'bulk' style work, the Watch Folder script seemed like a good place to start, but...
Changing workflows at Leeds DL

... our process for adding the digital objects to the repository is changing from an almost linear:

Request -> Digitise -> Enhance (e.g. crop) -> Save -> Deposit -> Catalogue

to something more like:

Request -> Digitise -> Enhance -> Save
\rightarrow Catalogue (including checking copyright) -> Deposit
Roles for repositories

- Large content collections (MOre content)
- Open access papers
- Research outputs and data
- Metadata aggregators and QA
- Preservation and grey lit (Finch)
- ...
Roles for repositories: the new world

- End of the great JISC repository era
- Friend: “second-class status being given to repositories by Government and other Establishment agencies in the UK”
- Finch-RCUK preference for gold open access
- PLoS ONE, PeerJ preprints …
- Bjork: “slowing growth” of subject repositories
- Vs ResearchGate, academia.edu?
Roles for repositories: workflow

- Workflow
  - Workflow
  - Workflow

- Identify role and USP, then construct, review, test ________ to attract MOre deposit (at scale), and provide more user benefits, for less work
Useful contacts

• Twitter: @depositmo
• Blog http://blog.soton.ac.uk/depositmo/
• Web downloads http://www.eprints.org/depositmo/
• EPrints Bazaar http://bazaar.eprints.org/
• Email sh94r@ecs.soton.ac.uk