



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?

DepositMOre

DepositMO

Impenetrable
acronym –
changing the

Modus Operandi of
repository deposit

June 2010 – Jan
2012

MO

to

DepositMOre

Self-explanatory goal
– getting more stuff
into your repository

Sept 2012 –

~~Mar~~ June 2013

MOre

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



Why SWORD?

Simple Web service Offering Repository Deposit

- SWORD (2007) single interface, deposit in multiple repositories. 'Fire and forget'.
- 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

Deposit an item - Step 3: Item details

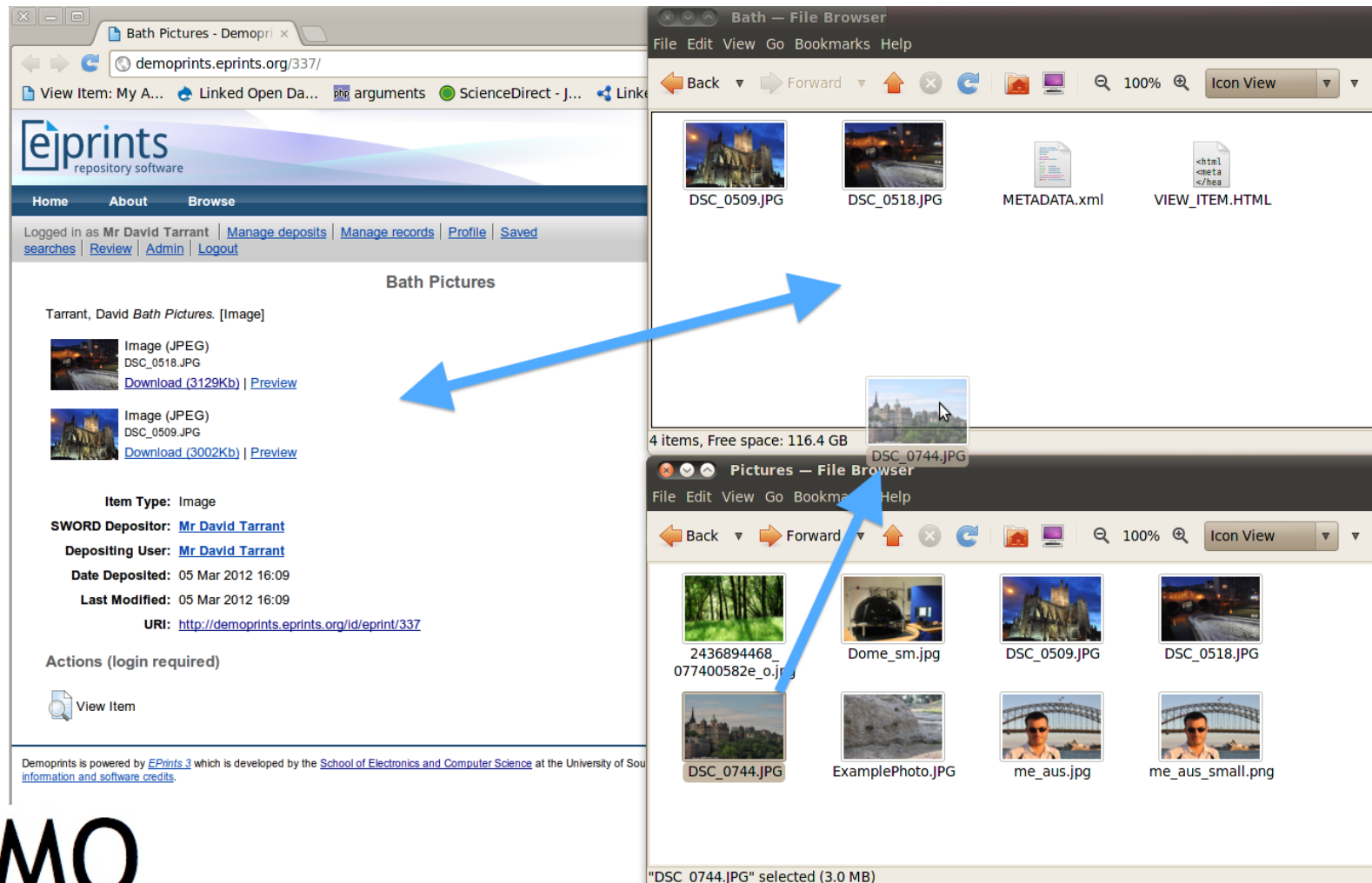
- **Deposit to:** <http://dspace.swordapp.org/sword/deposit/123456789/6>
- **Collection title:** Data sets
- **Abstract:** A collection for depositing data sets

Type of item:	<input type="text" value="Journal article"/>
Has the item been peer reviewed:	<input type="text" value="Yes"/>
Title:	<input type="text"/>
Abstract:	<input type="text"/>
1st author first name:	<input type="text"/>
1st author surname:	<input type="text"/>
2nd author first name (optional):	<input type="text"/>
2nd author surname (optional):	<input type="text"/>
3rd author first name (optional):	<input type="text"/>
3rd author surname (optional):	<input type="text"/>
Bibliographic citation (optional):	<input type="text"/>
Existing URL for item (optional):	<input type="text"/>

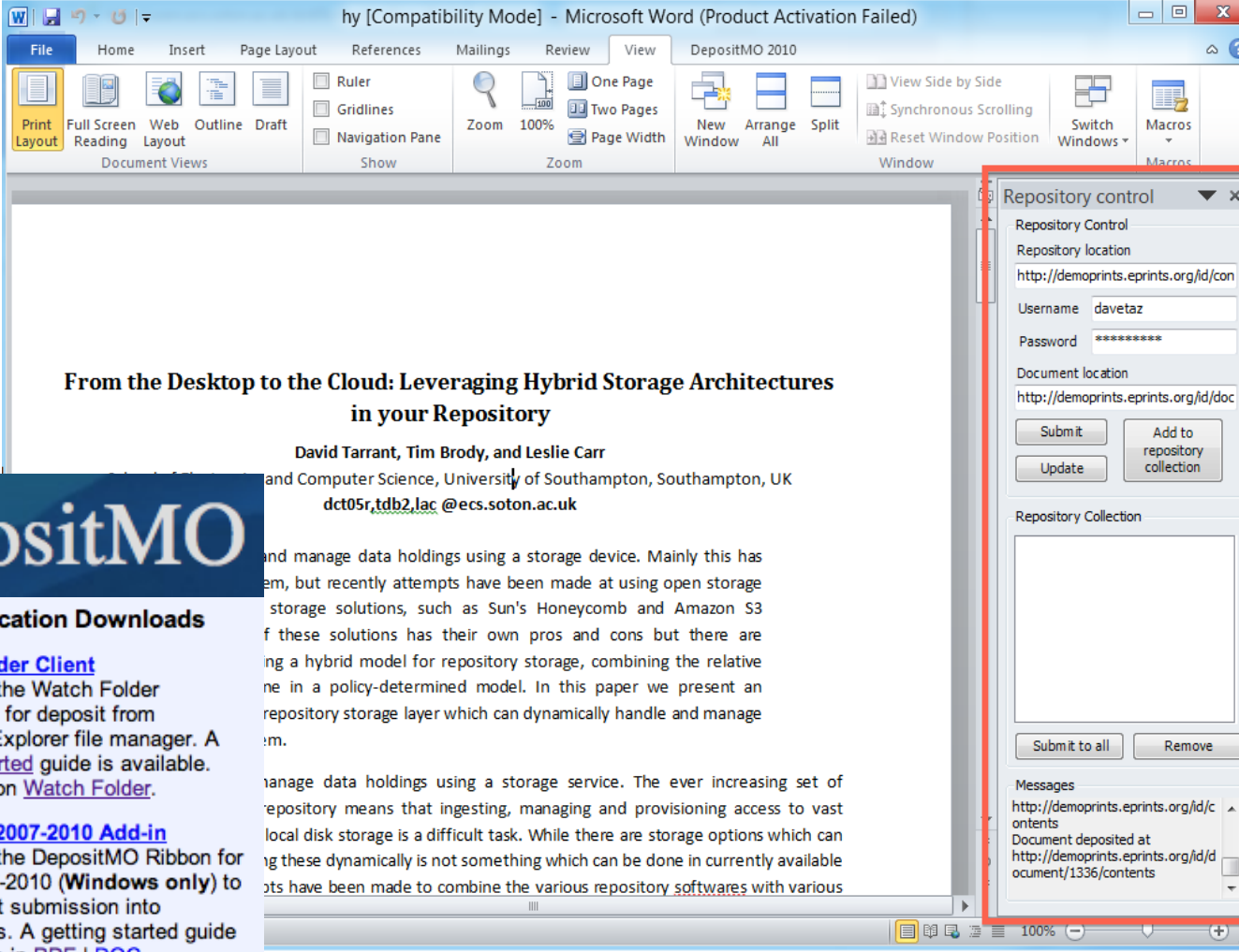
Facebook SWORD client

<http://blog.stuartlewis.com/2009/06/02/how-does-the-facebook-sword-client-actually-work/>

Watch Folder – a file manager-based deposit interface



Word Add-in deposit client



hy [Compatibility Mode] - Microsoft Word (Product Activation Failed)

File Home Insert Page Layout References Mailings Review View DepositMO 2010

Print Layout Full Screen Reading Web Layout Outline Draft Ruler Gridlines Navigation Pane Show Zoom 100% One Page Two Pages Page Width New Window Arrange All Split View Side by Side Synchronous Scrolling Reset Window Position Switch Windows Macros

From the Desktop to the Cloud: Leveraging Hybrid Storage Architectures in your Repository

David Tarrant, Tim Brody, and Leslie Carr

and Computer Science, University of Southampton, Southampton, UK
dct05r,tdb2,lac@ecs.soton.ac.uk

and manage data holdings using a storage device. Mainly this has been done using local disk storage, but recently attempts have been made at using open storage solutions, such as Sun's Honeycomb and Amazon S3. Each of these solutions has their own pros and cons but there are advantages to using a hybrid model for repository storage, combining the relative strengths of each in a policy-determined model. In this paper we present an open source repository storage layer which can dynamically handle and manage data holdings using a storage service. The ever increasing set of storage options means that ingesting, managing and provisioning access to vast amounts of data is a difficult task. While there are storage options which can handle these dynamically is not something which can be done in currently available software. Attempts have been made to combine the various repository softwares with various

DepositMO

Client Application Downloads

Watch Folder Client
Download the Watch Folder application for deposit from Windows Explorer file manager. A getting started guide is available. Blog post on [Watch Folder](#).

MS Word 2007-2010 Add-in
Download the DepositMO Ribbon for Word 2007-2010 (Windows only) to allow direct submission into repositories. A getting started guide is available in [PDF](#) | [DOC](#). Blog post on [Word Add-in](#).

Repository control

Repository Control

Repository location
<http://demoprints.eprints.org/id/con>

Username

Password

Document location
<http://demoprints.eprints.org/id/doc>

Repository Collection

Messages

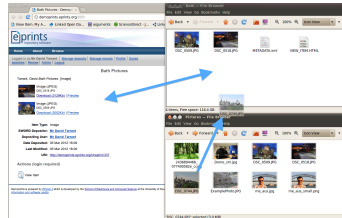
<http://demoprints.eprints.org/id/contents>

Document deposited at
<http://demoprints.eprints.org/id/document/1336/contents>

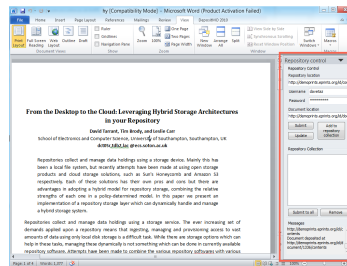
<http://www.eprints.org/depositmo/>

DepositMO/re tools

MO MOre



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.

X New

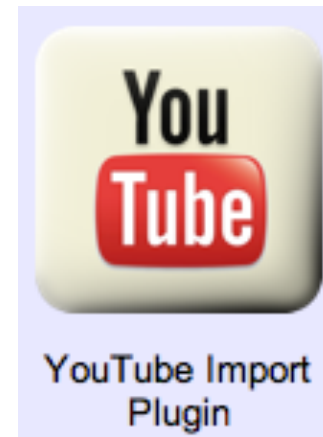
Enhancing and testing repository deposit interfaces, OR12, <http://eprints.soton.ac.uk/340997/>

Choice of content targets

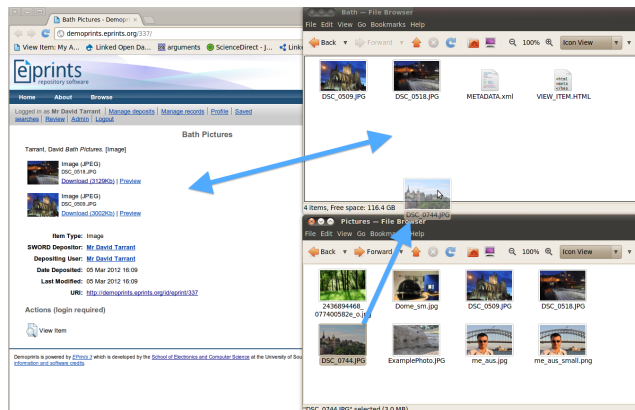
- EasyChair – conference papers
- Microsoft Academic Search
- ~~Europe PubMed Central~~ (RJ Broker – EDINA)
- arXiv
- YouTube/Vimeo – video
- Archaeology image generators (PTM)

MORe repositories: chosen tools

UCA
Goldsmiths
Glasgow



Leeds



Southampton (archaeology)



Bazaar Storerepository

Latest Additions



Tweepository



Data Cite DOI
Registration



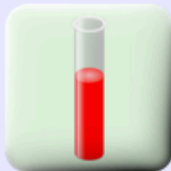
CORE Widget



Polynomial Texture Map
Extension



Arkivum A-Stor Storage
Backend



Labs/Testing Skin



Resource Manager



YouTube Import
Plugin



ReCollect



Eprints SOLR Integration



EPrints REF2014
Plugin



Encapsulate



Shelves for EPrints



Wordle Generator



Export plugin for CERIF 1.4
XML


5 June 2013

Bazaar Store repository



Apps by **MORE**

Southampton University on YouTube

 GB


University of Southampton

Q

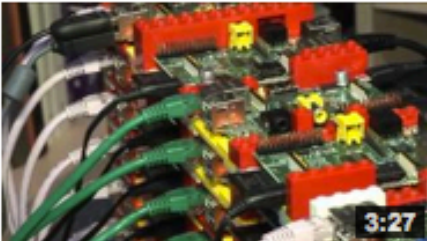
Upload ▼

Filters ▼


About 64,200 results



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.

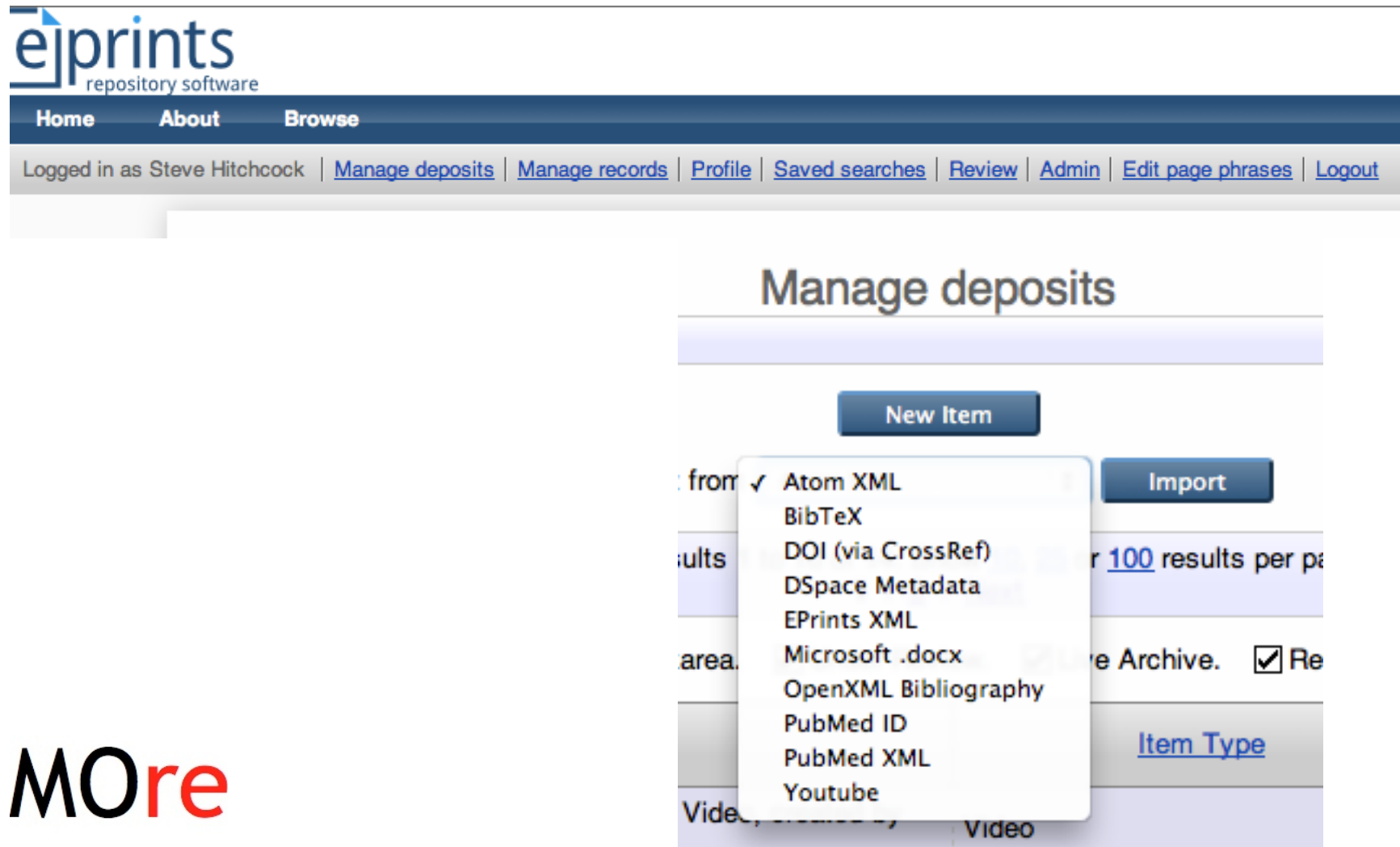


University of Southampton - Southampton engineers a Raspberry Pi Supercomputer
by **sotoncomms** • 7 months ago • 368,496 views
Computational Engineers at the **University of Southampton** have built a supercomputer from 64 Raspberry Pi computers and Lego ...



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

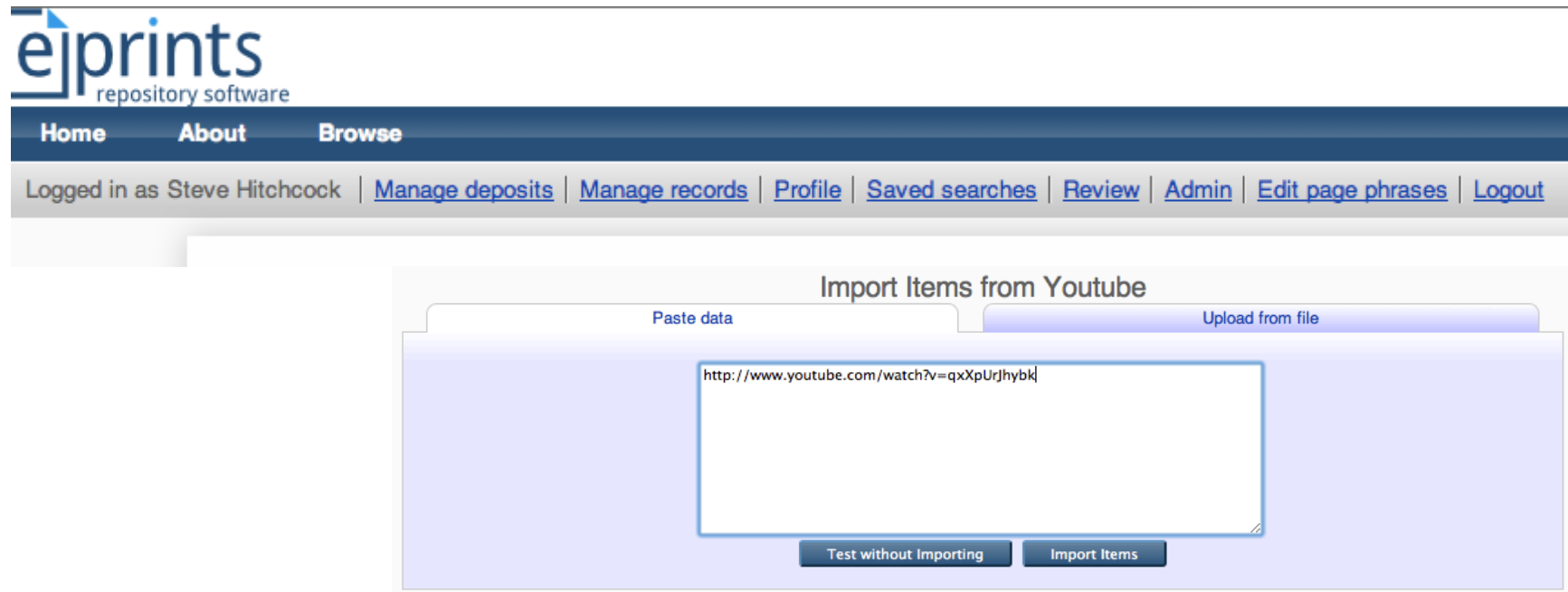


The screenshot displays the EPrints repository software interface. At the top, the 'eprints repository software' logo is visible. Below it, a navigation bar includes links for 'Home', 'About', and 'Browse'. A status bar indicates the user is logged in as 'Steve Hitchcock' and provides links to 'Manage deposits', 'Manage records', 'Profile', 'Saved searches', 'Review', 'Admin', 'Edit page phrases', and 'Logout'.

The main content area is titled 'Manage deposits'. It features a 'New Item' button, which has a dropdown menu open. The dropdown menu lists the following options: Atom XML, BibTeX, DOI (via CrossRef), DSpace Metadata, EPrints XML, Microsoft .docx, OpenXML Bibliography, PubMed ID, PubMed XML, and Youtube. To the right of the dropdown is an 'Import' button. Below these buttons, there is a search results section showing '100 results per page' and a checkbox for 'Re'. At the bottom, there is a table with columns for 'Item Type' and 'Video'.

MOre

Importing YouTube video: paste URL



The screenshot shows the 'ejprints repository software' interface. The top navigation bar includes links for 'Home', 'About', and 'Browse'. Below this, a user is logged in as 'Steve Hitchcock' with various management links. The main content area is titled 'Import Items from Youtube' and features two tabs: 'Paste data' (selected) and 'Upload from file'. The 'Paste data' tab contains a large text input field with the URL 'http://www.youtube.com/watch?v=qxXpUrJhybk'. At the bottom of the form are two buttons: 'Test without Importing' and 'Import Items'.

ejprints
repository software

Home About Browse

Logged in as Steve Hitchcock | [Manage deposits](#) | [Manage records](#) | [Profile](#) | [Saved searches](#) | [Review](#) | [Admin](#) | [Edit page phrases](#) | [Logout](#)

Import Items from Youtube

Paste data Upload from file

Test without Importing Import Items

MORE


Importing YouTube video: item type auto-identified

ejprints
repository software

[Home](#) [About](#) [Browse](#)


Logged in as Steve Hitchcock | [Manage deposits](#) | [Manage records](#) | [Profile](#) | [Saved searches](#) | [Review](#) | [Admin](#) | [Edit page phrases](#) | [Logout](#)

Edit item: University of Southampton - Southampton engineers a Raspberry Pi Supercomputer

 Import completed: 1 item(s) imported.

Type → Upload → Details → Subjects → Deposit

[Save and Return](#) [Cancel](#) [Next >](#)

★ **Item Type** 

☐ **Article**
An article in a journal, magazine, newspaper. Not necessarily peer-reviewed. May be an electronic-only medium, such as an online journal or news website.

☐ **Image**
A digital photograph or visual image.

☒ **Video**
A digital video.

☐ **Audio**
A sound recording.

MORE

Importing YouTube video: simple metadata auto-added

ejprints
repository software

[Home](#) [About](#) [Browse](#)

Logged in as Steve Hitchcock | [Manage deposits](#) | [Manage records](#) | [Profile](#) | [Saved searches](#) | [Review](#) | [Admin](#) | [Edit page phrases](#) | [Logout](#)

Edit item: University of Southampton - Southampton engineers a Raspberry Pi Supercomputer

Type → Upload → **Details** → Subjects → Deposit

< Previous Save and Return Cancel Next >

★ **Title** ?

University of Southampton – Southampton engineers a Raspberry Pi Supercomputer

Abstract ?

Computational Engineers at the University of Southampton have built a supercomputer from 64 Raspberry Pi computers and Lego. The team, led by Professor Simon...

★ **Creators** ?

	Family Name	Given Name / Initials	Email
1.	sotoncomms		http://www.youtube.c
2.			

! • Name in field [Creators](#) is missing the "given names/initials" part.

MOre

Importing YouTube video: done

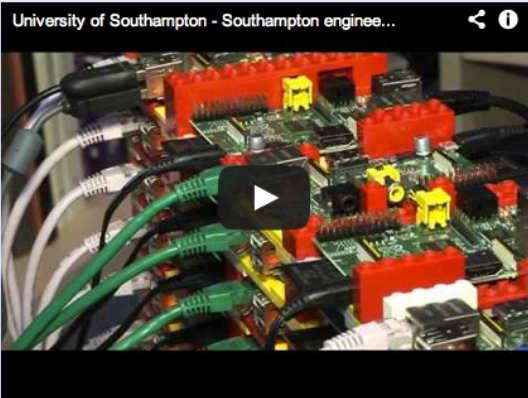
eprints
repository software


Home About Browse

Logged in as Steve Hitchcock | [Manage deposits](#) | [Manage records](#) | [Profile](#) | [Saved searches](#) | [Review](#) | [Admin](#) | [Edit page phrases](#) | [Logout](#)

Preview Details Actions History Issues

sotoncomms, sotoncomms (2012) *University of Southampton - Southampton engineers a Raspberry Pi Supercomputer.* [Video]

University of Southampton - Southampton engine... 

 Video
hqdefault.jpg

Restricted to Repository staff only
[Download \(22Kb\)](#)

Official URL: <http://www.youtube.com/watch?v=Jq5nrHz9I94>

Abstract

Computational Engineers at the University of Southampton have built a supercomputer from 64 Raspberry Pi computers and Lego. The team, led by Professor Simon...

Item Type: Video

MOre

Vimeo video deposited in GRO

RESEARCH ONLINE

Goldsmiths
UNIVERSITY OF LONDON

[Login](#)

[Search](#)

[Advanced Search](#)

Don't drum for other girls

[+ Tools](#)

Busby, Lisa, Harries, John, Hearn, Paul, Saez, Carlos, McClure, Bruce, Tovey, Ray and Horne, Christopher. *Don't drum for other girls*. [Audio]

Video (external)

Images

Videos/Audio Files

Information



[Home](#)
[About](#)
[Browse](#)
[Search](#)
[Help](#)
[Deposit Guide](#)
[Latest Additions](#)
[Policies](#)
[Statistics](#)
[Contact](#)

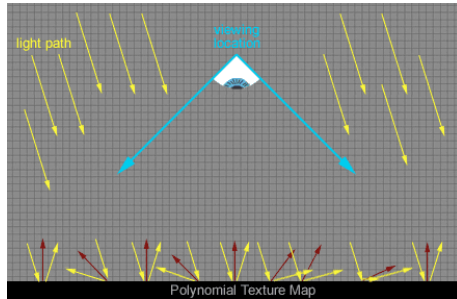
<http://eprints.gold.ac.uk/8284/>

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)



<http://www.flickr.com/photos/chough/3969335698/>



- Reflectance Transformation Imaging (RTI)
<http://culturalheritageimaging.org/Technologies/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/>

Unzipping a PTM archive: unaided



EPrints 3.3 test server – no PTM app

MOre

Unzipping a PTM archive: with a research data app and the PTM app

Manage deposits

[? Help](#)

New Item

Deposit as

New Dataset

Import from BibTeX

Import

EPrints 3.3 test server – with ReCollect data deposit app and PTM app

Add a new document

[?](#)

File

From URL

Choose File no file selected



Archive (ZIP)
4Mb



MORE

Spotting PTM unzip tool

The image shows a web interface for uploading a file to a PTM (Patent Theses Manager) system. The main form is titled "Archive (ZIP) 4Mb" and contains several fields for metadata. A red arrow points from a "PTM" folder icon in the top right of the form to a modal dialog box in the bottom right.

Form Fields:

- Title:** Text input field.
- Content:** Dropdown menu with "UNSPECIFIED" selected.
- Type:** Dropdown menu with "Archive (ZIP)" selected.
- Description:** Text input field.
- Visible to:** Dropdown menu with "Anyone" selected.
- License:** Dropdown menu with "UNSPECIFIED" selected. Other options include "The things license" and "UoS Thesis Copyright Coveragepage".
- Embargo date:** Fields for Year, Month (Unspecified), and Day (?).
- Embargo Reasons:** Dropdown menu with options: "Commercial Contract", "Patent Pending", "Publication Pending", "Ethical Confidentiality", "Third Party Copyright", "Ongoing Project", and "Other".









Modal Dialog:








The modal dialog is titled "EPrints User Area" and contains the following text and buttons:







- Archive (ZIP) 4Mb
- Unpack the contents of a PTM project directory .zip
- Buttons: "Unpack PTM" and "Cancel"







MOre




Unzipped: 3 files (it could be 80!)

	Archive (ZIP) 4Mb	     
		Hide options 
★ Title: <input type="text"/>		

	Other (Polynomial Texture Map configuration) 22Kb	    
		Hide options 

	Other (Polynomial Texture Map (PTM)) 1702Kb	   
		Hide options 

	Image (Source Images) 3378Kb	   
		Hide options 

★ Dataset Information		
★ Title:	<input type="text"/>	
★ Type of Data:	<input type="text"/>	
	<input type="text"/>	
Description/Abstract:	<input type="text"/>	

MORE

Users' views on EPrints PTM app

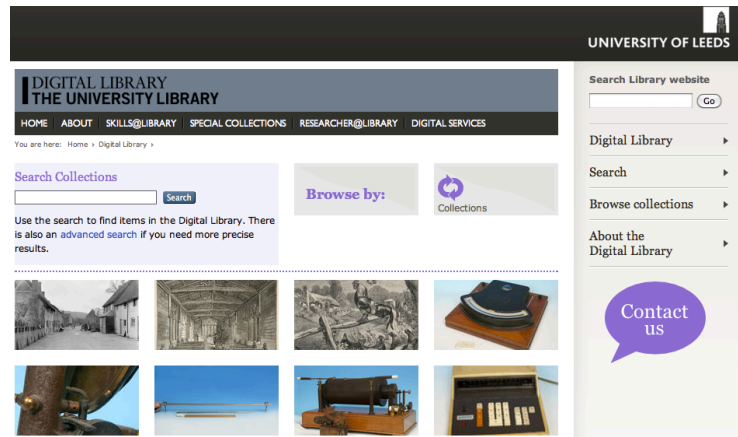
- The system is almost there. There are some usability issues, but I guess most of them are related with the eprints itself, not the ptm app. The final unzipped content page needs a bit of work, as at the moment you cannot access original files one by one, only by downloading whole zip.
- I think the 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.

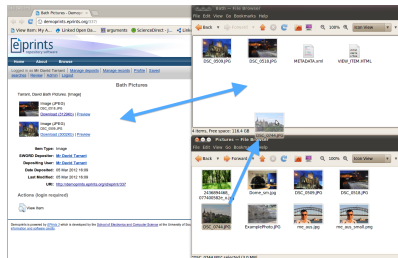
Graeme Earl, Southampton archaeologist

New Digital Library at Leeds



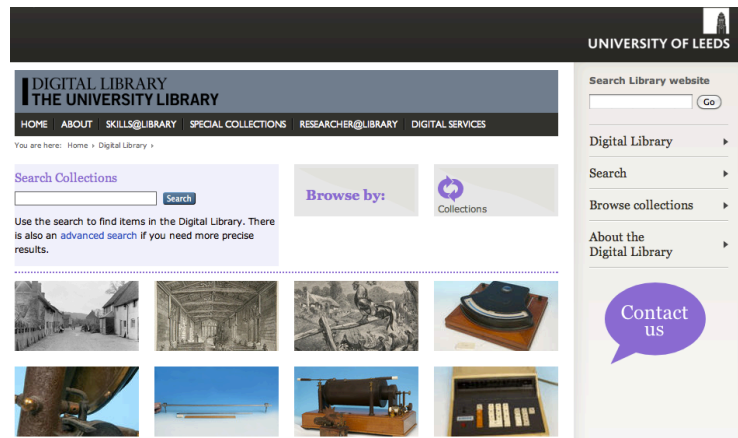
Based on the EPrints platform ...

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...



MO

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 -----\
 \-> 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