**Enhancing and testing repository deposit interfaces**

Steve Hitchcock, David Tarrant and Les Carr

School of Electronics and Computer Science, University of Southampton, UK

sh94r, davetaz, lac @ecs.soton.ac.uk

**Abstract**: Native content deposit interfaces of popular repository softwares such as DSpace and EPrints have changed little since their introduction over a decade ago. The more recent emergence of SWORD, a means of abstracting the deposit interface from the underlying repository software, provides opportunities for experimenting with new deposit methods. The Developer Challenge event at recent Open Repositories conferences has provided award-winning examples [1,2]. The JISC DepositMO Project has developed two new repository deposit interfaces, in the process both contributing to and using the new SWORDv2 protocol. These interfaces aimed to reposition deposit in the context of popular authoring and desktop management tools, also to reschedule repository deposit from end-of-creation to an in-progress activity, thereby seeking to change the culture, or *modus operandi* (the ‘MO’ of DepositMO), of repository deposit. These interfaces have been extensively tested with users. This paper illustrates use of these interfaces and summarises the test results. We discovered it is possible to demonstrate an initial wow! factor for repository deposit, but this is harder to sustain in practice. Briefly, we describe a short follow-on proposal, which aims to adapt and apply these tools to demonstrate increased rates of deposit of repository content.

**1 Introduction**

Digital repositories essentially provide a series of interfaces to get things done, such as deposit content in the repository, manage content (usually an administrator, but can involve others) and access content (users can browse or search, or machines can harvest content using OAI-PMH). It’s curious, then, that there have been few published studies or user testing of repository interfaces.

In DepositMO (June 2010 to January 2012) we were concerned with one of the repository interfaces, for deposit – that is, how new content is added to the repository. Both EPrints and DSpace provide native deposit interfaces, configurable by repository administrators; another popular repository software, Fedora, requires third-party interfaces. Basically, these deposit interfaces consist of Web forms that collect information, or metadata, describing the item or digital object being deposited, and a button to start the download of the selected item from a specified location.

These forms have been criticised for being too long and taking too much time to complete, even if the claim may be shown to be exaggerated (3). There is a perennial trade-off between collecting sufficient information from the depositor of an object at the point of deposit to ensure it can be fully described and searchable, and minimising inconvenience and time taken by the depositor. Thus for digital repositories, user-facing issues in supporting deposit include design and requirements for metadata as well as more general Web usability.

**2 SWORD**

The idea that repository deposit might be abstracted from repository software began in practice with SWORD (Simple Web service Offering Repository Deposit) in 2007. At the core of its proposition was, given the plethora of repositories and user interfaces, that users could choose a single interface for deposit in multiple repositories. SWORDv2 takes it from a ‘fire and forget’ deposit method that is limited in terms of what authors can do with their content after depositing in a repository, to a model where authors retain more control and items can be created, updated, replaced, or deleted (CRUD) in the repository.

SWORDv2 is the basis of DepositMO’s vision for in-progress deposit. Through DepositMO, SWORDv2 is available with current versions of EPrints (3.3) and DSpace (1.8). In addition, the aim of bringing more powerful content management features to repositories led DepositMO to specify two key extensions to SWORDv2, which are similarly included in the current DSpace and EPrints implementations [4]:

**3 New repository deposit clients**

To take advantage this new deposit support in repository software and demonstrate deposit of in-progress digital works, DepositMO built deposit tools into familiar authoring and desktop applications: Microsoft Word, and a file manager.

**3.1 Word Add-in**

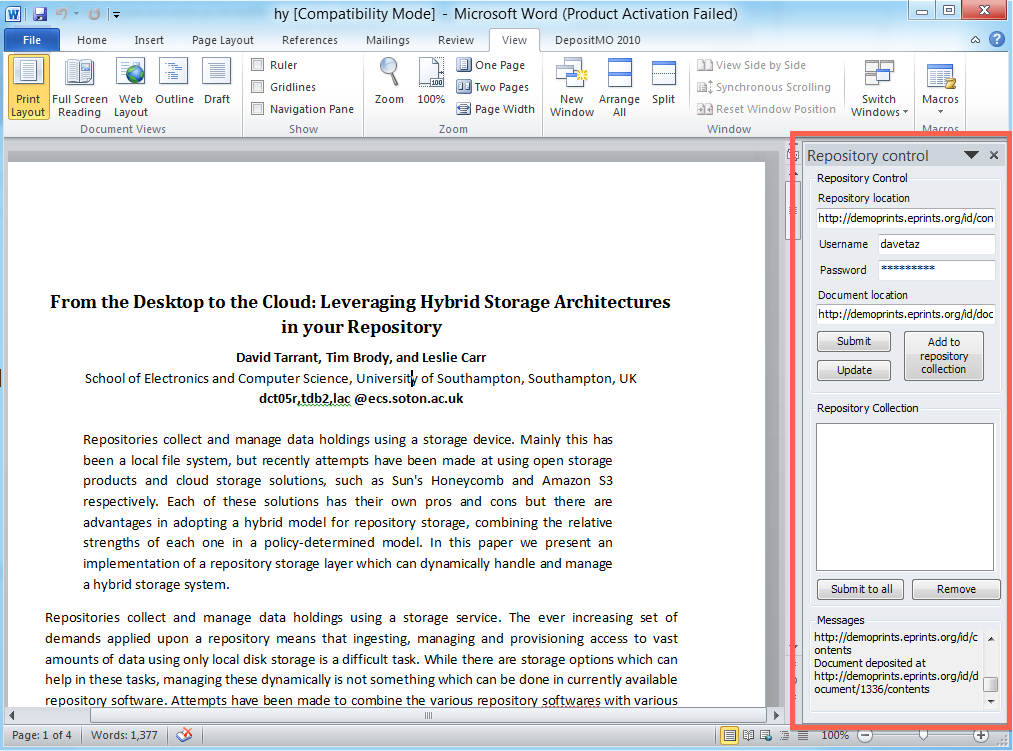
The Word Add-in produced by DepositMO allows content in a document to be saved to, updated, or deleted from a repository during composition, without leaving the application, and is available to download and install with Microsoft Office/Word 2010. We chose to work with Word as it is widely used among researchers and scientists to write research papers and create materials that are likely to be deposited in institutional repositories (5). In addition, Microsoft has been seeking to extend the use of Word for digital science by supporting plugins, applications that provide additional functionality required by science authors, such as science and maths markup and DTDs. The deposit tool described here is another such plugin.

Figure 1. Word showing instance of repository deposit panel

The Add-in is accessed via ribbon tabs in Word, and displays as a repository control panel in the Word window (Fig. 1). The panel allows details of the target repository – repository location and user login – to be entered. Documents can be submitted to the repository and updated by clicking a series of buttons provided, and the progress of submissions is displayed as messages in the panel.

**3.2 Watch Folder – a file manager-based deposit interface**

Computer systems provide users with file manager tools, such as Windows Explorer or Mac Finder. Users can save files to designated directories and folders, and can move and copy files to different locations using short-cut keys or drag-and-drop.

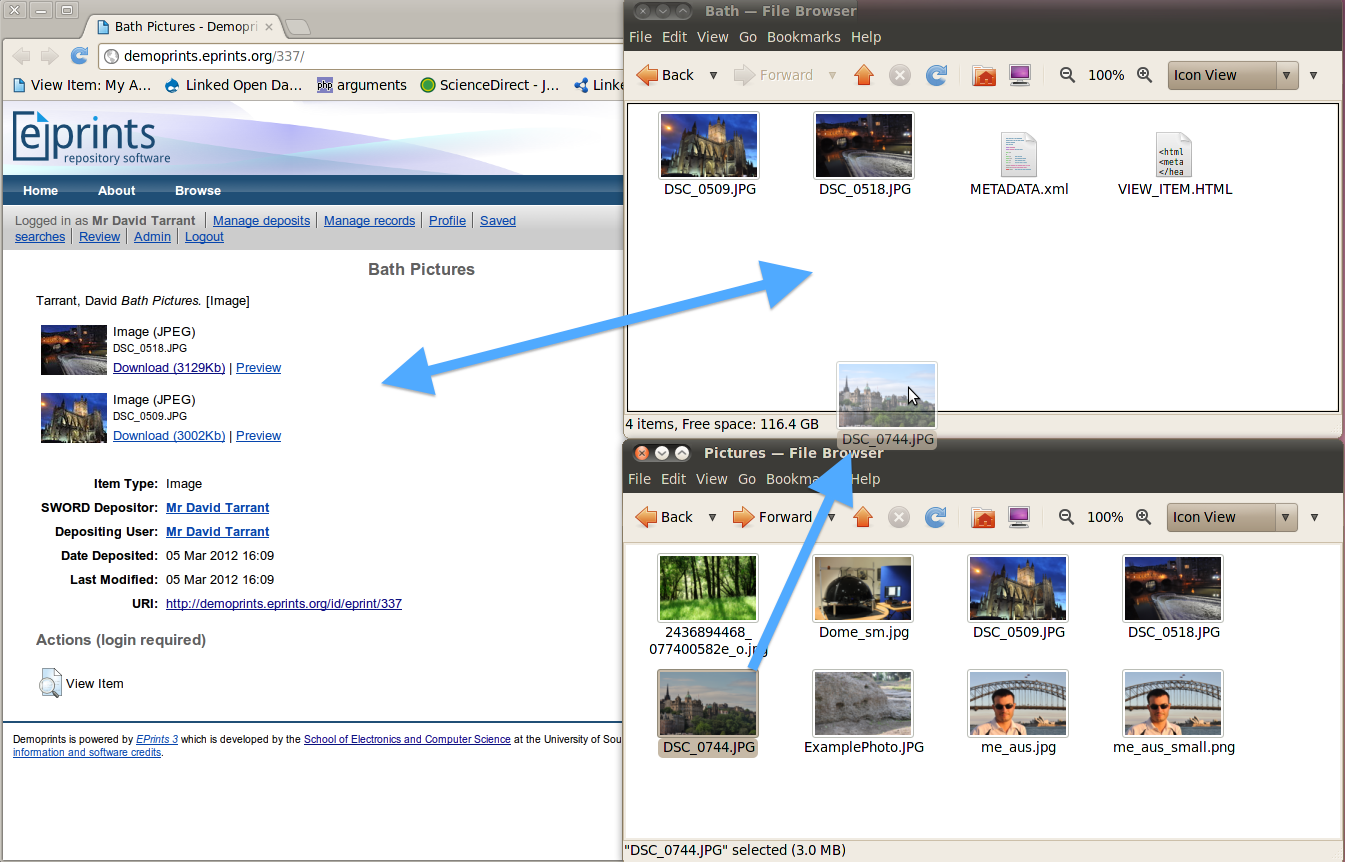
Using this metaphor for content management, DepositMO makes use of drag-and-drop by working with file manager tools, creating a DropBox-style client to initiate repository deposit. It works on the basis of a downloadable ‘watch folder’ script. When a new file is dragged to the directory or folder being ‘watched’ the script initiates a session with the selected repository and copies the file there as well (Fig. 2). The view of the target repository is activated by clicking on VIEW\_ITEM.HTML in the watched folder. Direction to the repository – repository URL and user login, again – is provided to the script via a simple text form launched by clicking on the METADATA.xml file in the same directory.

Figure 2 In a file manager application an image from a user’s collection (lower right) is dragged into a ‘watched’ folder on the user’s hard drive (upper right) to be replicated in the target repository (left)

In this scheme, a folder in the active directory containing the watch folder represents a record in the repository, and a file in that folder represents an object linked from that record. Thus, folders can be used to organise collections in the repository.

**4 Summary findings from user tests of new deposit clients**

When first demonstrated live these tools received spontaneous applause [6]. To test the process of repository deposit using these two new tools more formally, they were set up in a controlled test environment: a pair of Web-connected laptops running Windows 7, directed at a demonstrator EPrints repository running the SWORDv2 extensions. Test users included new as well as experienced repository users, since the tools are intended to open new deposit scenarios, with most working in observed pairs. Results were based on what users did, using repository records and observer notes; what users said, in before- and after-test questions; as well as task completion and time taken. The full results are available in a series of posts on the project blog [7].

Results from these tests suggest that, on average, both direct deposit clients (Word Add-in and Watch Folder) took less time to deposit an item than via the native repository interface. The principal difference between the new tools and the native interface is in providing detailed metadata: there is a time advantage where less stringent metadata requirements apply or where the tool provides some assistance with metadata completion, as the Word Add-in does with simple bibliographic information, for example. In its present form the Watch Folder does not provide the simplest means of metadata control, and in some cases additional time was incurred to complete this stage.

The wow! factor seen when these tools were first demonstrated is harder to sustain in practice, noticeably as a user’s collection size grows and issues of metadata control and versioning become more critical. It’s not just about the initial deposit. There are no simple comparisons to be made between the instant deposit tools (Word Add-in and Watch Folder) and the more structured native repository deposit interfaces. Instant deposit may be at the expense of providing careful metadata now: further refinements to the tools might be able to improve metadata control without losing deposit time.

**5 Impact on real repositories – DepositMOre**

Over 50% of test users indicated these new deposit tools would encourage them to submit more of their own content to the repository, but these are usability tests and cannot tell us how these tools might impact on deposit rates of real repositories. A short project, DepositMOre, was proposed to demonstrate increased deposit rates using in-tool statistical tracking. The project adopted the content managing interface – the Watch Folder, and a derived interface to locate and deposit a user’s existing Web content – rather than the interface linked to a content creation tool like Word, where impact on deposit rates will be less immediate.

DepositMO recognised that repositories must extend their user interaction model for intelligent bulk ingest to deal with dozens or even 100s of documents at once. DepositMOre re-examined the Watch Folder for this purpose with some of the original user testers from within arts and archaeology, two communities that generate large volumes of non-textual digital objects and need efficient services for storing and accessing this content, a role that could be performed by repositories.

Another potential source of repository content is research papers from online conference management services. The EasyChair Deposit Tool used in DepositMOre lists a user’s authored items in Easychair, which hosts over 15,000 conferences. The tool also checks if these items are present in the user’s selected repository. Any that have not been deposited can be added with one click.

**6 Conclusion**

DepositMO went some way towards simplifying the process of submitting desktop-locked content into a repository. When merging a private environment with a public one the clear challenge for any client is to be simple, concise and easy to use.

DepositMO and DepositMOre show how different repository deposit tools can support users with different deposit demands, widening the base of repository users. With the growing emphasis on managing research data, especially using data repositories, the need for choice in repository deposit – offering tradeoffs between time to deposit and degree of documentation – is going to become more acute.

**References**

1 FedoraFS by Rebecca Koeser, RepoChallenge Winners! [dev8D](http://dev8d.jiscinvolve.org/wp/) blog, May 20, 2009 <http://dev8d.jiscinvolve.org/wp/2009/05/20/repochallenge-winners/>

2 Repository as a Service (RaaS), Glimpse into the Future of Repositories: videos now available! JISC [Digital Infrastructure Team](http://infteam.jiscinvolve.org/wp/) blog, August 4, 2011 <http://infteam.jiscinvolve.org/wp/2011/08/04/glimpse-into-the-future-of-repositories-videos-now-available/>

3 Carr, L. and Harnad, S., Keystroke Economy: A Study of the Time and Effort Involved in Self-Archiving. Technical Report, ECS, University of Southampton <http://eprints.ecs.soton.ac.uk/10688/>

4 DepositMO Profile: SWORDv2 Extensions <http://swordapp.org/docs/DepositMO.html>

5 Tarrant, D., Carr, L., Wade, A. and Warner, S., Interactive Multi-Submission Deposit Workflows for Desktop Applications. Open Repositories 2010, July, Madrid <http://eprints.ecs.soton.ac.uk/20839/>

6 Dave Tarrant demos MS Kinect & SWORD v2 deposit, Repository Fringe 2011, Edinburgh, ‪repofringe YouTube Channel, Sep 16, 2011 <http://www.youtube.com/watch?v=kPZ3n6sJqaU>

7 User testing results, Modus Operandi for Repository Deposits blog, February 3, 14, 2012 <http://blogs.ecs.soton.ac.uk/depositmo/tag/user-testing-results/>