

Another look at desktop repositories

This post is a refresh on where we are at with the idea of Desktop Repositories, sparked by a discussion I had this week with people from [Intersect](#), the NSW eResearch organisation where they are working on the next version of [FieldHelper](#), tool that originated at The University of Sydney. I met with Anne Cregan, Chris Kenward and Raul Carrizo on Monday 2011-02-14¹.

The idea of FieldHelper is to allow field workers (archaeologists, linguists, anthropologists etc) to manage and describe their data and get it off their hard drives in to 'staging repositories'. The new round of work on FieldHelper will be funded under the Australian National Data Service (ANDS) data capture program. I wanted to make sure that the Intersect team were aware of the work we've been doing towards similar aims, and to let them know about the software we've built. My team at USQ is funded by ANDS to work on software for managing metadata about research data, the [ReDBox application](#). This work is under the ANDS metadata stores program, and is therefore focussed on describing captive data rather than the process of capture, but my team has had an interest in data mustering for a while, and we have done some small pilot projects. My thinking is that if there is enough of a match, then our work might give the new FieldHelper a bit of a head-start. If not, that's useful data too.

First, some history. I [proposed the idea of a desktop eResearch repository](#) in 2009, inspired by FieldHelper and by metadata-aware consumer tools for media-data management like Picasa and iTunes. I was also inspired by some comments from Mr Eprints himself, Les Carr. This was picked up by a few people and I [talked to people in Sydney about the ideas](#). I posted [an early view of what the architecture might look like](#).

In summary, the idea was to provide a platform for researchers that would:

1. **Watch their hard disk** (on the laptop, desktop, fileshare or instrument) for new files.
2. Show them the files in a web browser – emphasising that even though the files are still private, and local **the web is everything**.
3. Provide a **plug-in framework for handling any kind of file**, extracting technical metadata, indexing it to death, and making it web-viewable.
4. Allow the researcher to **organise and classify** the files using both formal and informal metadata and send collections of stuff off to be backed up, preserved or disseminated.

The ADFI team, in a project led by Duncan Dickinson did some work with a researcher at USQ, Leonie Jones, on a Vietnam War history project exploring what a media-rich desktop repository for history research might look like, and demonstrating how some of the content she's working with can be pushed up to the World Wide (as opposed to desktop) Web for a community of veterans to interact with the materials Leonie has collected; tagging and geolocating photos, identifying people in pictures using the official roll held at the Australian War Memorial, and reconstructing the official history. This is covered in a couple of conference presentations from eResearch Australasia [1]. And you can [buy the DVD](#). We didn't help make the DVD, mind, but we would like to help manage the resources that went into making it and future productions.

All this work has been on hold at ADFI, as we work on the ReDBox application for research metadata, but we have restarted some of it in the wake of [Beyond the PDF](#). Very soon we'll launch [a little site which allows people to put stuff in Dropbox™ and then see it with in a few minutes on the web](#). Oops – looks like that was a launch, of sorts. That [link](#) is to a package I made out of the sample files from the BTPDF workshop, which [Anita de Waard](#) refers to as “The Bourne Corpus”. All I did was drop them in Dropbox, wait a minute or two for them to sync and then I was able to browse to the files thru the web interface, select them, click to package the selected files, and then drag-n-drop them into the order I wanted them.

So, the package contains a bunch of stuff including some draft versions of the document, meaning it is a bit of a mess, but it does show how we can publish stuff like the [supporting data](#) – or [info about the data](#) or [stuff](#)

I wonder what it was about their building, which seems to house a lot of insurance companies. There were lots of couriers delivering roses, and blokes carrying bunches of same. The place reeked of romance. Strange.

[about the grant](#). Log in with any old OpenId and you can make your own packages. If you'd like to try the system on your own files, drop me a line and I will share the Dropbox with you. The demo will probably be around until mid 2011 – these links are not guaranteed persistent. There are bugs. Batteries not included. If you want your own, stay tuned and I'll post when we have installation instructions.

Copyright Peter Sefton, 2010. Licensed under Creative Commons Attribution-Share Alike 2.5 Australia.
<<http://creativecommons.org/licenses/by-sa/2.5/au/>>



This post was written in OpenOffice.org, using templates and tools provided by the [Integrated Content Environment](#) project and published to WordPress using [The Fascinator](#).