More on negative click or net benefit repositories

So the conversation that Chris Rusbridge started about low-effort repositories rolls on. Chris summarizes some of the responses. Including mine and broadens the discussion to bring in some of the stuff that Andy Powell has been saying:

> Andy wants repositories to be more consistent with the web architecture. He spoke at a Talis workshop recently; his slides are here (on Slideshare, one of his models for a repository).

This reminded me that earlier this year people in my network were talking about Andy's keynote at VALA. We responded to the ripples running through the Oz-repos community by putting a project proposal to ARROW in Australia to start working on a repository ingest application that is much more 'of the web' than those we have now.

The ARROW board didn't approve that one, I'm sure it wasn't the just the name that was wrong but I gather that was not popular. And it was a truly stupid name.

I had to think of something quickly so I called it VICE-SQUAD in the spirit of highly contrived acronyms that seems to pervade the ARROW community.

> VICE SQUAD means (VITAL-compatible Integrated Content Environment-driven Service-oriented Queryable User-friendly Application for Data-acquisition)

Here's a bit of the proposal we put, which seems to be along the lines of what Chris from the Logical Operator blog suggested in response to Chris R. From our proposal:

> The goal of this project is to build a smart user-friendly repository ingest system for VITAL and/or other Fedora based repositories, which will be implemented in the Integrated Content Environment (ICE) service framework. The system will be released as open source software. The application will be a stand-alone ingest system with back-end coupling to ICE.
>
> The project will attempt to create an innovative interface for repository ingest which is quite different from other approaches, allowing users to upload content into a w*orking repository*, or *workbench* from where it can be shared with the world (sharing with defined groups is out of scope for this project but will be dealt with in a separate USQ project) and/or submitted for ingest into the repository; ie pushed over the *curation boundary*[1].
>
> It will consist of three interfaces:
>
> 1. **A dead-simple user interface** for academics to share their work as quickly as possible and tag it with free-form metadata. They will upload items to a workbench where they will be able to work on them further, or merely mark them for ingest into the repository.
>
>   (see this blog post from the JISC repositories interest group for some thinking along the same lines, with pointers to a commercial service called box.net which could serve as a model for the sharing-features proposed here if adapted to an academic context.)
> 2. **A graphical user interface for repository staff** or advanced users to edit MODS

metadata for a record and turn the user's initial tags into formal metadata, including the ability to edit existing metadata records from VITAL.

3. A **seamless** tie-in to a structured authoring environment, so that papers authored in such an environment can be sent to a repository with a single click

In addition to the two interfaces there will be behind-the-scenes 'smarts' that can extract metadata from documents and produce HTML and PDF automatically, using technologies already developed by USQ.

I think the time has come for someone to build a repository which has the simple ePrints approach to collecting metadata, with an option to make it even simpler and just go with tags if that's all the energy the depositor can muster.

Our proposal goes on to talk about MODS and MARC and METS but I think maybe the time is right to do RDF, especially if the Bibliographic Ontology makes it into Zotero. And we should look at ORE support rather than bother with METS.

For those who care to add more higher-quality metadata – and often this a librarian tidying up later – there needs to be just a little bit more smarts than ePrints or DSpace offer with their flat metadata in the area of stuff like research affiliation and researcher identity, stored in RDF, with an option to serialize it in other metadata formats as required.

While we didn't get that project up with ARROW we will have another opportunity to build on the forthcoming TheOREM-ICE work.

We have a big need for simple sharing in ICE right now, and I imagine that this will be true for thesis writing too – wouldn't it be great to share your PhD draft with reviewers and draft-readers in a simple way?

One thing I'd like to do is to turn on document sharing via an obscure non-guessable URL so that people can drop in and comment on my documents using ICE's inline annotation systems without authentication. Or for more formal collaboration, I want to be able to create ad hoc workgroups preferably using a single sign on service of some kind. Once we get through some of the nasty issues we're having with the ICE 2 beta version we will no doubt start adding those collaborative features.

Then when TheOREM kicks off we'll have an ICE to repository gateway pumping content into DSpace, Fedora and ePrints.

What's needed as well are some simple services to let people upload stuff and push it out. ICE already lets you push to a blog via ATOM (all the posts here are done that way), but we could add SlideShare and Flickr and suchlike as additional services, as well as a simple web sharing interface that is less controlled than the Institutional Repository. As Peter Murray-Rust says: Don't use Institutional Repositories put it on the web.