Crossing curation mountain

I'm looking forward to seeing Peter Murray Rust [eat my dog food](). He's lucky cos at our place the hounds eat relatively benign dry food. Oh, and raw chicken wings and mouldy bread and diverse leftovers deemed to foul for [the fowl](). Not to mention compost, fertilizer of all kinds of origin, grass and lizards and parrots when they can get them. Maybe not so lucky, but I'll buy him a beer to wash it down.

I'm going to follow up on using ICE for blogging to WordPress soon which is what that dog food stuff is about, but Peter has just pointed out some issues with getting papers into institutional repositories and I wanted to discuss some of his points here.

In summary I think that his message is that the **curation boundary is too high** to cross for authoring teams. Or is it that it is **too far away** from the authoring process?

Here's his summary:

> [Using our own repository]()
>
> ...
>
> Here are some basic problems about reposting:
> - the process from starting a manuscript to final publication can take months or years
> - there are likely to be multiple authors
> - authors will appear and disappear during the process
> - manuscripts may fission or fuse.
> - authors may come from different institutions

PMR then goes on to talk about the solution they're trying. They have set up an authoring repository when collaborators use Subversion (SVN):

> Well, we are starting to write our papers using our own repository. Not an IR, but an SVN repository. So Nick, Joe and I will share versions of our manuscripts in the WWMM SVN repository. Joe wrote his thesis with SVN/TeX and I think Nick's doing the same. Joe thought it was a great way to do things.

([ICE]() uses Subversion too and it works well for us, but we have had to put quite a complex layer in front of it so ordinary non-technical people can deal with it. We have been asked to support LaTeX; I'll let you all know if that comes off.)

> The advantage of SVN is that you have a complete version history. The disadvantage is only that it's not easy to run between institutions. I am not a supporter of certificates. And remember that not all our authors are part of the higher education system. In fact Google documents starts to look attractive (though the versioning is not as nice as SVN.)
>
> Will it work? I don't know. Probably not 100% - we often get bitten by access permissions, forgetting where things are, etc. But it's worth a try.

Two things here:

1. Working between institutions is definitely a problem, particularly when not everyone is in higher-ed. I'm not sure what kind of certificates PMR's talking about but I will say this: [Shibboleth]() isn't going to solve the problem when there are groups coming from all over the place any time soon, but [OpenID]() might help. Using OpenID the owner of a working repository should be able to tell SVN who to trust, using their OpenID – and if you come from an institution that doesn't use it then you can always get one from elsewhere.

   Should work for web application, but it could be a challenge to get the commandline version of SVN

working with OpenID because some OpenID providers might require you to use a web-GUI to authenticate, but people have thought about it.

2. Google docs? Well as discussed here before Stijn Dekeyser and Richard Watson of USQ found that it makes a good online text editor with supernatural collaborative powers. If you want to edit in raw LaTeX that is. They made a demo system that lets you render Google docs, even. But as a word processor? No. You'd be stopped by the lack of support for stuff like bibliographies and Maths before you got too far and don't get me restarted on the HTML it makes.

I liked the last bit, so I added some emphasis:

And **if I were funding repositories I would certainly put resource into communal authoring environments**. If you do that, then it really is a one-click reposition instead of the half-day mess of trying to find the lost documents.

I'll be sure to mention this to our friends at DEST.