

Deflation in repository clicks

Chris Rusbridge is contributing to repository click deflation at the Digital Curation Blog with a post about [Negative Click Repositories](#).

Why deflation?

At Open Repositories 2008 a group of us Australian developers entered in the Repositories Challenge, with an entry entitled [Zero Click Ingest](#) [1]. The introduction puts it like this:

This micro-project demonstrates a way to eliminate the repository deposit step altogether, by having the repository software take responsibility for collecting the content that it needs. It involves using the Integrated Content Environment¹ (Sefton 2006)[2] (ICE) as a document authoring system, but the principle could be applied to other content management systems which support metadata or category-aware ATOM or RSS feeds, with the ability to supply the requisite formats. We show how documents created and managed in ICE can be automatically ingested into a repository at the appropriate time, based on document state.

So we did **zero click** but now we have to go for negative clicks? That's not fair!

There are two aspects of this I'd like to comment on.

1. Chris is calling for other people's take on the negative click repository, so I'll contribute my thoughts.
2. It would be wrong to think that we can only get people to use new software if it means they don't have to learn anything new. I don't buy that; people can change when they see the benefits. The trick is to get them to want to change by themselves.

My thoughts on repository workflow

I've been thinking about repository workflow ever since I joined the [RUBRIC](#) project in 2005. I was outraged that this Institutional Repository business seemed to be all about putting PDFs into ugly web sites. Here were all these custodians of collections of a few thousand PDF files getting together at conferences to talk about how they were doing Web two-dot-oh. Um, no you're not. Try using a web format first. The program committee at Open Repositories 2007 didn't accept my paper along those lines but they let me give a [poster](#). That poster is available in HTML, PDF and as a slideshow, all generated from a single source document. Now that's at least web 1.01.

In 2006, I talked about some of the possibilities for smart content management systems (like ICE) to supply content directly to repositories, in [Workflow 2.0](#).

We still haven't done most of what I talked about in Workflow 2.0, although the Zero Click Ingest demo was a big step in proving the concept. But we're about to make a big push in this direction in collaboration with Jim Downing and team at Cambridge as part of the [JISC TheOREM project](#).

In TheOREM we're going to set up ICE as a 'Thesis Management System' where a candidate

can work on a thesis which is a true mashup of data and document aka a datument [3]. When it's done and the candidate is rubber-stamped with a big PhD, the Thesis management system will flag that, and the thesis will flow off to the relevant IR and subject repositories, as a fully-fledged part of the semantic web, thanks the embedded semantics and links to data.

One more idea: consider repository ingest via Zotero or EndNote. In writing this post I just went to the OR08 repository and used Zotero to grab our paper from there so I could cite it here. It would be cool to push it to the USQ ePrints with a single click.

That's one click but it's about fifty **less** than the alternative. Is that what you mean by negative, Chris?

Net-benefit repository workflows

I'm not sure I'm happy with the implications of the term *negative click* here (even 'zero clicks' is a bit suss). As Chris puts it:

So we need to develop repositories that make the data work to take human work away: negative click repositories.

Not that I don't agree that we want to take the work away, just that I think this will involve change and I think we should be careful not to imply that we can perform magic.

At USQ, the [Integrated Content Environment](#) is on the way to becoming a 'core' system for producing courseware. It has been available for a few years under an open source license, but as far as I know it's not used much at all outside of USQ. Why?

I think its because it represents a net loss for most users; they don't need to do what ICE does and/or they can't afford the time to set it all up. This is why almost nobody has taken it up spontaneously for general writing tasks or even for courseware outside of our institution.

But at USQ, where we are reaffirming our commitment to flexible delivery – we call it Fleximode – staff know they **have to** create resources that suit on-campus, web and print use. ICE helps with that, so it has grown organically from our first user to a couple of hundred because overall it makes life easier. Learning ICE is not trivial, you have to do some training, you have to change the way you work, and the organization needs to supply support. If you do use it, though, there's a net benefit.

I think it is a mistake to focus on the Negative Click repository as a goal and assume that it's going to happen without changing user behavior. If people are bashing away in word processors as though they are typewriters then you miss the opportunity to create XHTML reliably, and you throw out heaps of semantics. We'll be stuck with PDF repositories for a long time until we empower our authors to do better.

If we can seed the change by empowering some researchers to perform much better – to more easily write up their research, to begin to embed data and visualizations, to create semantic-web-ready documents – then their colleagues will notice and make the change too. That's two things:

1. Make it easier.

2. Make the result sexier.

This is one reason why TheOREM is so exciting; not that it's going to look at ORE but that it will be a first step to providing tools for PhD candidates and their supervisors that I hope will be the envy of others, just as Shirley Reushle used the first version of ICE to make an online course that met the USQ standards and her colleagues saw it and wanted to do the same.

At one point in the development of ICE we were beating off potential users. I reckon we can do the same with eResearch and repositories but it's hard work and it takes time.

- [1] L. Monus et al., "Zero Click Ingest," Apr. 2008; <http://pubs.or08.ecs.soton.ac.uk/119/>.
- [2] P. Sefton, "The Integrated Content Environment for Research and Scholarship," *ICE Website*, 2006; http://ice.usq.edu.au/introduction/ice_rs.htm.
- [3] P. Murray-Rust and H.S. Rzepa, "The Next Big Thing: From Hypermedia to Datuments," *Journal of Digital Information*, vol. 5, 2004, p. 248; <http://jodi.tamu.edu/Articles/v05/i01/Murray-Rust/?printable=1>.