

I've never been much of a note-taker. At uni I tended to use the subliminal absorption method in lectures. That is, sit there, while doodling designs for cycle panniers or tents, or writing the name of a certain young woman over and over in as close as I could manage to Helvetica font and try to stay awake after finishing my taxi shift at 3am and hope that something sinks in even if I do doze off a bit. I'm not apologizing. Seems to have worked so far for me.

So, I'm not equipped with the notes or the disposition to be able to give a [blow-by-blow of the ETD conference](#) the way RUBRIC's Neil Godfrey has done, but I will offer my impressions of the two ETD (Electronic Theses and Dissertations) conferences that I have attended.

ETD 2005

I attended the ETD conference in Sydney in 2005, with Deidre Lowe, but apparently I didn't blog anything, probably because we were waiting for confirmation of the [RUBRIC](#) project to come through and we were keeping it low-key. My recollection of that conference is that it was a good introduction not only to the idea of electronic theses, but also to the world of repositories. And I got large, but not fatal dose of Open Access evangelism, because Stevan Harnad was there enthusiastically and persuasively [Maximising research impact by mandating institutional self-archiving](#).

But the thing that struck me most was that just about everyone in 2005 seemed to be resigned to putting PDF into their repositories. There was a group from Humboldt University in Germany talking about XML, and a couple of other mentions for XML, but none of them had worked out how to put authoring tools into the hands of large numbers of authors and automate the publishing process. You can look in the references to [my paper for ETD 2007](#) for further reading.

ETD 2007

I was itching to get back to this community and show them what we can achieve with ICE – which I did at ETD 2007.

But what about the rest of ETD 2007?

The thing that I picked up, mainly from the plenary sessions, is that there is an increasing amount of attention being paid to e-scholarship, or e-research.

[Neil Godfrey blogged](#) about Greg Crane's presentation on e-scholarship in the humanities:

Greg Crane spoke of the need and inevitability of moving beyond book-imitation pdf files. He used [Peseus Classics Online](#) as an example of the potential we should be aiming towards — where texts contain multiple links for each word — to dictionaries, to other related texts, to commentaries. The potential impact will move us beyond the slow and limited intake of information that comes currently from reading lines at a time, then moving on to other texts a 2 dimensional process as opposed to the 3 dimensional or more organic structure possible with the sort of thing we now see at Perseus.

I don't know the technical structure behind Perseus, but I know Perseus well enough to see it as one model for a future online database — and as for metadata implications, what it is calling for is work on ontologies and the semantic web (i suspect perseus is not based on that at present but i could be wrong — and I see Greg has an article online discussing this Perseus project in more depth that I must read) — and that means RDF ideally rather than traditional schema such as MODS or MARC or DC. — though the RDF structured content could generate such schema when needed. (My thoughts arising from Greg's presentation.)

<http://metallogger.wordpress.com/2007/03/28/thesis-types-in-repositories/>

There was also Peter Murray Rust's [tour of some collisions between e-science and publishing](#).

From my point of view there were more issues raised in the e-scholarship that questions answered. These are my questions. Things I'm thinking actively thinking about and experimenting with at the moment:

- What is the relationship between research data, classical texts, source materials etc and publications derived from those data?
- Which kinds of data should go into institutional repositories? Which kinds should not?
- How do we identify and refer to all the bits and pieces involved over time as they are created, edited, published revised and archived?
- What's the relationship between print and online hypertextual delivery?
- How can we help research communities to manage their data for access and preservation? (They don't necessarily do it all that well by themselves, but it can be hard to tell people that)
- How can we help researchers to choose and configure tools?

A lot of the issues look like the sorts of things that we deal with in the ICE project; instead of dumping Microsoft word on the desktop, like a typical ICT department, and leaving people to suffer, we can help them to use it productively so they get more out of their time and their work.

In contrast to some e-Research meetings where things tend to be discussed a policy level only, without real-life examples, one of the good things about the ETD conference is that it features lots of down to earth basic implementation stories, including a set of awards for ETDs.

For example, there was presentation on how to embed video in a PDF file. I didn't say anything at the time, but I think that this is a particularly bad idea – it's not likely to be usable even in the short term because it is so fragile regarding codecs and platforms and over time it will have huge preservation problems, not to mention usability. I'd prefer to see this addressed along the lines of the way I tackled in in my paper, which looked at how multimedia objects might be managed with appropriate renditions for print and online versions of the work.

But the point is that people are there presenting stuff about real things they have tried, however modest, which is a great strength for the conference series.

