Webb's Web

More on my blog www.alia.org.au/webbsblog

WEBB'S WEB

Opening up the data

A recent online discussion has caused me to think more about the Open Data movement, especially the segment where government agencies are being urged to release their data resources so that the community can make use of them. The problem is that the community (by and large) has shown remarkably little interest in the concept, even after a series of contests to create apps to exploit the datasets.

The discussion featured a fair bit of pontification about the cleanliness and integrity of the data, the lack of metadata, the relatively small number of datasets and so on. It reminds me of the time when I worked in CSIRO and our team of researchers has some real neat software that would enable people to search through a video library and edit a number of clips to produce a custom video – perhaps the greatest tries of their favourite Rugby football player. The problem then was that the potential users weren't interested in doing that – they'd look at something that someone else had produced but they just didn't want to put the effort into doing it themselves. So it is, I believe, with open data.

If a bunch of enthusiasts put together an app that merges bus timetables and maps and sells it for a very small price, then people may be interested. But as for doing it themselves: no. They have lives to lead.

A further consideration is how to get the data together in a state to be published. The conversation I mentioned above did raise some good points. There is definitely data available, but it was never meant to see the light of day; it was just intended to be a by-product of other services. No wonder it's poorly presented and badly described (if described at all). And to bring it up to publication quality takes money, and that's hard to come by. No amount of pious boosting by people in central organisations will overcome the reluctance to find that extra money by the workers who are custodians of the data (and most of them don't even realise that they are custodians).

So spare us the hand-waving, and if you're serious about the topic then why not open up the purse strings and fund the data provision. And while you're about it, make sure that it's high on the list of priorities of the services that the community really wants you to spend their money on.

The future of the book – and it can't come fast enough

If there's one thing we love more than books, it's talking about them – and I don't mean the explosion in book clubs or in contributing reviews to Amazon and similar sites. No, this time my object is the (apparently) never-ending discourse on what's going to take the place of the familiar paper object that we want to touch, smell, hear and finally devour. This train of thought originated in two essays: the first was on audiobooks at **nplusonemag.com/listening-to-books** by Maggie Gram. In this she lists what works well (for her) in this medium and what doesn't, and then looks at the history and current usage of audiobooks, including one which appeals to many of my friends who suffer through long commutes (although I know that over-engaging with a book as your drive can be a recipe for disaster). In the second essay by Tim Parks (**tinyurl.com/76pfmzv**) one of his interesting proposals is that the abstraction from the object that you get with an eBook brings you closer to the essence of the author's intent. I think that when the newness of these alternatives wears off in a few years we'll read all of these ramblings with a great deal of head-shaking amusement.

Fix it so it doesn't break

It's been more than a year since he gave his talk to the Royal Society titled *Web Science: A New Frontier* but Jonathan Zittrain's words are still as relevant as they were all those months ago. His paper at **papers.ssrn.com/sol3/papers.cfm?abstract_id=1995059** is currently published on the SSRN site, but registration is free and relatively easy. He's generally speaking about different ways in which the web might "break" and he explores issues like international copyright restrictions, URL shorteners, archives and suggestions for new forms of caching. It's worth reading.

There's not a moment to lose!

I've been thinking for quite a while about the impermanence of web resources and what we can do about it. That's why I was very interested to see this story about the Archive Team (**www.technologyreview.com/web/39317/**). They're a sort of squad that leaps into action when they get word that a service is going to be shut down – like Yahoo's Geocities or Poetry.com – and sets about grabbing as much of the publicly accessible material as they can. Sometimes their rescue mission is not welcomed by the hosts, but the rest of us can only applaud their guerrilla tactics.

Posted by Kerry Webb