

# Webb's web

## A little warning

You may not have seen it but Google does us all a favour whenever they find a site containing malware, by putting a note in their search results that "This site may harm your computer". In a post to their official blog <http://tinyurl.com/64gswm>, they point out how this might come about and what you as a webmaster can do about it, if you're unlucky enough to be compromised. They also discuss whether or not they should inform the webmaster before they flag the site in your search results, and make a convincing case that the users' rights should prevail. I still marvel at the processing power that they have at their disposal to do this sort of detection.

## Patron on the line

There's a service called Mosio <http://www.mosio.com> where users send SMS messages to ask questions of the community (before you rush to try it, the bad news is that it's not available here yet). Looking at their website, some of the questions are pretty basic: what is veal; how many cups in a pint; is infidel a word? Hmm. For slightly more difficult questions, they've introduced a new service called Text a Librarian <http://www.textalibrarian.com/> which takes a little time to understand but seems to involve your library paying for the service that your patrons can use to submit questions to reference staff. I'll be interested to see how it takes off.

## A set of useful statistics

As an old ABS hand, I've always believed that you can't have too many statistics (a view that appears not to be held by the current government). That's why I was so impressed by the report released by Opera <http://tinyurl.com/47hun4> and not only because it reinforced a few of my prejudices—although that's not really a bad thing. Some of the highlights are: around 4% of pages had valid HTML, around 33% had Flash content (this did vary considerably between countries) and 80% used CSS. Further on the HTML front: only 50% of those displaying 'valid HTML' badges were actually entitled to that claim.

## Metadata musings

I was at a meeting of the AGLS Working Group in October; it's a jolly little gathering that happens twice a year, run very effectively by Adrian Cunningham from the National Archives, and he pointed out that the AGLS metadata standard was just ten years old. Yes, back in September 1998, it came into being after a busy year of development. It all started off being based closely on the Dublin Core elements, but pretty much from the start there was a degree of divergence. At first, the AGLS meant Australian Government Locator Service (a passing reference to the GILS system from the US Government) and it was in the Commonwealth Government (with some interest in the states and territories) that it began to have an influence.

Within a couple of years though, the Working Group realised that there would be benefit in a wider adoption of the product,

and negotiations began with Standards Australia for AGLS (now known only by the initials) to become a standard for all Australians, whether in Government or not, and so we saw *AS 5044* emerge in 2002. And there it starts to get interesting, because I think it's reasonable to want to know if it's been all worthwhile. Is anyone outside government using it? Standards Australia has been asked about the number of copies of the Standard that have been sold, but apparently that's a closely guarded secret.

It could well be that no copies of this document have been sold at all - possibly because it's always been available for download from the AGLS part of the NAA website at <http://tinyurl.com/678cj9>. But I guess we'll never know.

But to return to the subject of how much it's used, we really do need to know this. In the beginning, there was a general expectation that use of the standard would result in a golden age of findability. Well it happened—but not in the way that we expected. The search engines (which in 1998 were not as pervasive or as powerful as they are now) soon lost interest in using META tags as a source of indexing material, when unscrupulous types began to put all sorts of irrelevant keywords in those tags. And of course, only the zealots actually took the time to put quality metadata on their sites. But at the same time, the search engines kept improving their indexing and search capabilities so that we can now get the sort of results we hoped for, just from the page content itself.

Metadata is still being used, but mostly in more specialised applications like service portals and places like HealthInsite <http://www.healthinsite.gov.au> that can afford professional cataloguers. I doubt if it will ever make a mark outside those specific applications. So it goes.

## Talking to themselves

Some of the luminaries in Web Issues are now saying that they won't accept any more comments on their posts, because it's too hard to manage them, and that's not what the blog is for, and so on and on. Fortunately, that not the case with many of the most useful blogs around (like Matt Cutts' at <http://www.mattcutts.com/blog/>) but it just shows me how much our little endeavours are still in their infancy.

Kerry Webb  
[kwebb@alianet.alia.org.au](mailto:kwebb@alianet.alia.org.au)

Don't forget to visit my blog  
<http://www.alia.org.au/webbsblog>

