

# The Trouble with Bibliographies

Martin Fenner, Gobbledygook

March 7, 2011

The bibliography of a scholarly paper is interesting and important reading material. You can see whether the authors have cited the relevant literature, and you often find references to interesting papers you didn't know about. Bibliographies are obviously also needed to count citations, and then do all kinds of useful and not so useful things with them.

Bibliography by quinn.anya at Flickr.

Unfortunately almost all bibliographies are in the wrong format. What you want is at least a direct link to the cited work using the DOI (if available), and a lot of journals do that. You don't want to have a link to PubMed using the PubMed ID as the only option (as in PubMed Central), as this requires a few more mouse clicks to get to the full-text article. And you don't want to go to an extra page, then use a link to search the PubMed database, and then use a few more mouse clicks to get to the full-text article (something that could happen to you with a PLoS journal).

A bibliography should really be made available in a downloadable format such as BibTeX. Unfortunately journal publishers – including Open Access publishers – in most cases don't see that they can provide a lot of value here without too much extra work. One of the few publishers offering this service is BioMed Central – feel free to mention other journals that do the same in the comments.

This weekend Peter Murray-Rust invited Peter Sefton and me to Cambridge (UK) for a very interesting workshop about **Scholarly HTML**. Our goal is to discuss how we can define standards and build tools to make HTML the best platform for scholars and scholarly works. The event is in fact a hackfest, and we hope to have something to show by Sunday evening.

My idea for the hackfest is a tool that extracts all links (references and weblinks) out of a HTML document (or URL) and creates a bibliography. The generated bibliography should be both in HTML (using the Citation Style Language ) and BibTeX formats, and should ideally also support the Citation Typing Ontology (CiTO) and COinS - a standard to embed bibliographic metadata in HTML. I will use PHP as a programming language and will try to build both a generic tool and something that can work as a WordPress plugin. Obviously I will not

start from scratch, but will reuse several already existing libraries. Any feedback or help for this project is much appreciated.

If I had a tool with which I could create my own bibliographies (and in the formats I want), I would no longer care so much about journals not offering this service. One big problem would still persist, and that is that most subscription journals wouldn't allow the redistribution of the bibliographies to their papers. A single citation can't have a copyright, but a compilation of citations can. I'm sure we will also discuss this topic at the workshop, as Peter Murray-Rust is one of the biggest proponents of Open Bibliographic Data.