

What is Digitometrics?

An interview with
Tim Brody
by **Sara Kjellberg**, Lund
University Libraries, Head
Office



Tim Brody, a PhD candidate in the Intelligence, Agents, Multimedia Group at the University of Southampton, has been involved in the Open Access movement for four years. My first question to Tim was: what exactly does the term Digitometrics mean? Tim was introduced to the work with Digitometrics in the Open Citation project, which he was part of in his undergraduate project. He tells me that Digitometrics is “bibliometrics for the online age”. You combine the results from citation analysis with web logs, counting the number of times an article has been downloaded, to get an article's impact.

Tim has set up a service based on the content in arXiv, the physics subject archive, to provide a citation analysis service called Citebase. He says that the arXiv data are quite simple and easy to process since the authors use a common citation style. “Although there are some problems to be aware of before we can see more solutions like Citebase in other subjects, e.g. all subjects have their own way of formatting their references. One big problem is that we still lack a lot of content in many subjects, and the archives provide little internal linking to base the analysis on”. What Tim means is that if a human being can understand a reference it is also possible to instruct a computer what to do with it in order to make sense of it.

The linking in the arXiv data has been solved because of the use of quite consistent references; and one can use the numbers for the volume, page number, and years to identify a reference. Tim tells me that you need something that is common to connect the references together, and he says that numbers are more reliable than for example the title text string. The Citation linking can be

used for example to add functionality like showing related articles.

I asked when we will see a more general service like Citebase for all subjects? Tim answers that we will probably see different services for different purposes in the new world of Open Access publishing. He sees the Open Access environment as divided into three layers: one layer with self-archived full texts and Open Access publishers e.g. arXiv; a second layer where you will find citation analysis and linking services like Citebase; and a third layer for information retrieval with search engines and navigation tools, for example Google. This does not mean, that he thinks that we will see a new science citation index-like database for Open Access publications in all subjects, but more likely that we will find the more generic services in the top layer for search engines followed by subject based services, or services divided some other way, based on data from the repositories in the other two layers.

One of the important aspects of Digitometrics will be that we change focus from rating journals to actually rating the articles. I find it very interesting that Tim investigates the possibility of using statistics to look for patterns and predict the future impact of an article. Tim is involved in a joint project to study the difference in citation impact for published Open Access articles versus articles available only through subscriptions. One could conclude from Tim's and his colleagues' work that we will probably see a lot happening in the near future in how we can measure the impact of an article. Tim rounds up our talk by saying, that he thinks that we are seeing a lot of momentum behind Open Access, and that we probably will see more solutions for different subject areas as we see the content grow in the archives.

References:

<http://opcit.eprints.org/>

<http://citebase.eprints.org/cgi-bin/search>

<http://www.arxiv.org/>

Brody, T., Kampa, S., Harnad, S., Carr, L. and Hitchcock, S. (2003) Digitometric Services for Open Archives Environments. In Proceedings of European Conference on Digital Libraries 2003 , pages pp. 207-220, Trondheim, Norway.

<http://eprints.ecs.soton.ac.uk/archive/00007503/>