

2006-10-16

Should Arbetslivsinstitutet be closed? Nordic Webometric Comparison

Ulf Sandström, Linköping University
&
Mike Thelwall, Wolverhampton University

Web links are a useful source of indicative information about the online impact of a web site and the context in which a web site is found useful. For example, the search engine Google's ranking algorithm is based upon the belief that it is possible to distinguish more useful from less useful pages and sites by counting the links that point to them (Brin & Page, 1998). It is reasonable to think of links to a web site as indicating its "online impact" (Ingwersen, 1998) or the trust that others have in it (Davenport and Cronin, 2000), perhaps paralleling the citation impact of the site's owners. In academia, links are not only used to point to high quality information, however, but also to important organisations and producers of information, as well as for other reasons (Wilkinson, Harries, Thelwall, & Price, 2003). Hence, link data has to be treated with caution because of the variety of reasons for which links are created as well as the lack of quality control online (Bar-Ilan, 2001; Björneborn and Ingwersen, 2001). The best use of link data is probably in a comparative role: comparing the online impact of similar web sites can highlight interesting differences and makes it easier to detect high and low inlink counts (Thelwall, 2004). It is also possible to use inlink counts to investigate the sphere of influence of an organisation: for example, are the links predominantly national or international?

Nordic Institutes in Occupational Health and Comparable Institutes

Data on research institutes was taken from the Swedish EUROLABS database covering research institutes in Denmark, Sweden and Finland.¹ Updates of the Nordic part of the database were done during 2005. For the comparative project reported in this paper we have chosen institutes of approximately the same size according to their financial resources FY 2000 and 2003.

Sweden:

<http://www.arbetslivsinstitutet.se/>

<http://www.stfi.se/>

<http://www.smittskyddsinstitutet.se/>

Arbetslivsinstitutet

STFI-Packforsk AB

Smittskyddsinstitutet

Denmark:

<http://www.ami.dk/>

<http://www.dmu.dk/>

<http://www.dfum.dk/>

<http://www.delta.dk/>

Arbejdsmiljøinstitutet

Danmarks Miljøundersøgelser

Danmarks Fiskeriundersøgelser

DELTA-tekniskt udviklingsforetag

¹ Information on the EUROLABS project is to be found in Appendix 1.

Finland

<http://www.ttl.fi/>
<http://www.ktl.fi/>

Finnish Institute of Occupational Health
 National Public Health Institute

Analysis of Nordic Institutes

The first test is of how many web pages link to each web site. The statistics in Figure 1 are the average of the results from MSN and Yahoo!. The linkdomain: query was used and links from the same site excluded. For example, for <http://www.stfi.se> the query used was linkdomain:stfi.se -site:stfi.se which should match all links to any page in the stfi.se web site, except pages from the same web site.

The statistics in Figure 1 could be called “web visibility”. Institutes that have attracted many links seem to be well-known online. Other studies have shown that this kind of web visibility correlates quite well with research output, so it would not be surprising if the biggest and most research productive institutes attracted the most links.

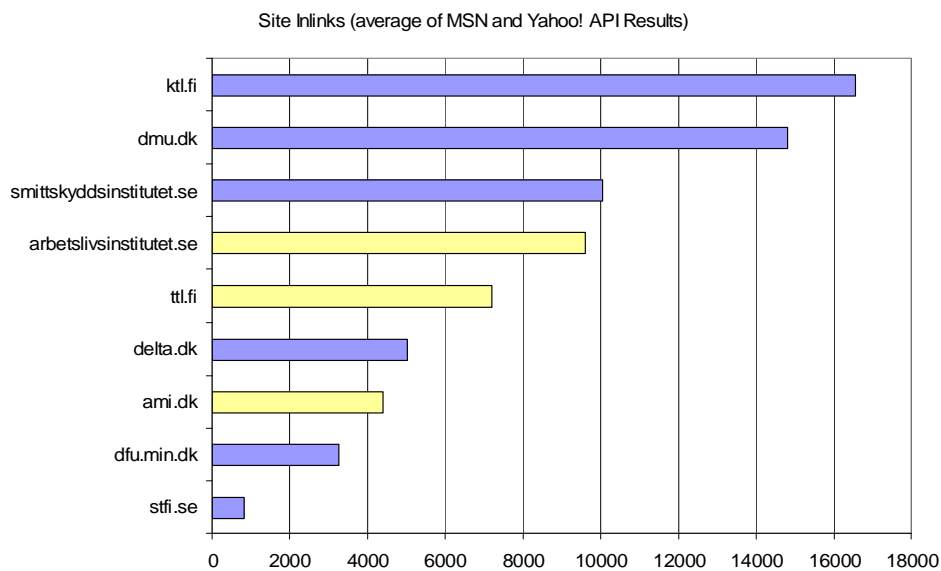


Figure 1. The web visibility of Nordic research institutes (Occupational health in yellow)

Breakdown of links

The second test is a breakdown of the main sources of links to each web site. It is based upon the top 999 pages linking to each web site reported by Yahoo! (999 is a technical limitation of Yahoo!). From the list of 999 link pages multiple pages from the same site were excluded to prevent individual sites contributing many links. Then the top-level domains were extracted from each remaining page URL and summarised in Figure 2. This could be called the web sphere of influence of each institute. Some institutes seem to be highly national in focus, whereas others are more international. The picture has to be interpreted cautiously because the generic domains .com, .net and .org contain a wide variety of types of information, from commercial to educational, and are widely used internationally so should not be read as US-based. Similarly, the

national domains are used for a wide variety of purposes although they are probably reliable indicators of national interest.

Arbetslivsinstitutet (National Institute for Working Life) in Sweden

Figure 2 shows the type of pages that link to the Nordic institutes. The results indicate that the Arbetslivsinstitutet attracts rather a lot of links from .com web sites. This is an indicator of relations and networks with commercial companies. The Arbetslivsinstitutet even has a larger number of such links than the industrial research institutes in this sample, both Swedish (STFI-Packforsk) and Danish (Delta). In addition, regarding relations to .se web sites the Arbetslivsinstitutet has a strong performance, but there are examples of organisations with more links, e.g. Smittskyddsinstitutet in Sweden and TTL (Finland), both in medical areas. Surprisingly, the Arbetslivsinstitutet has more international inlinks than the Smittskyddsinstitutet, which has been in the public domain to a large degree due to the bird flu hype during 2006.

Compared with AMI in Denmark and TTL in Finland results indicate that the Arbetslivsinstitutet is more linked from national and international organisations (.net and .org).

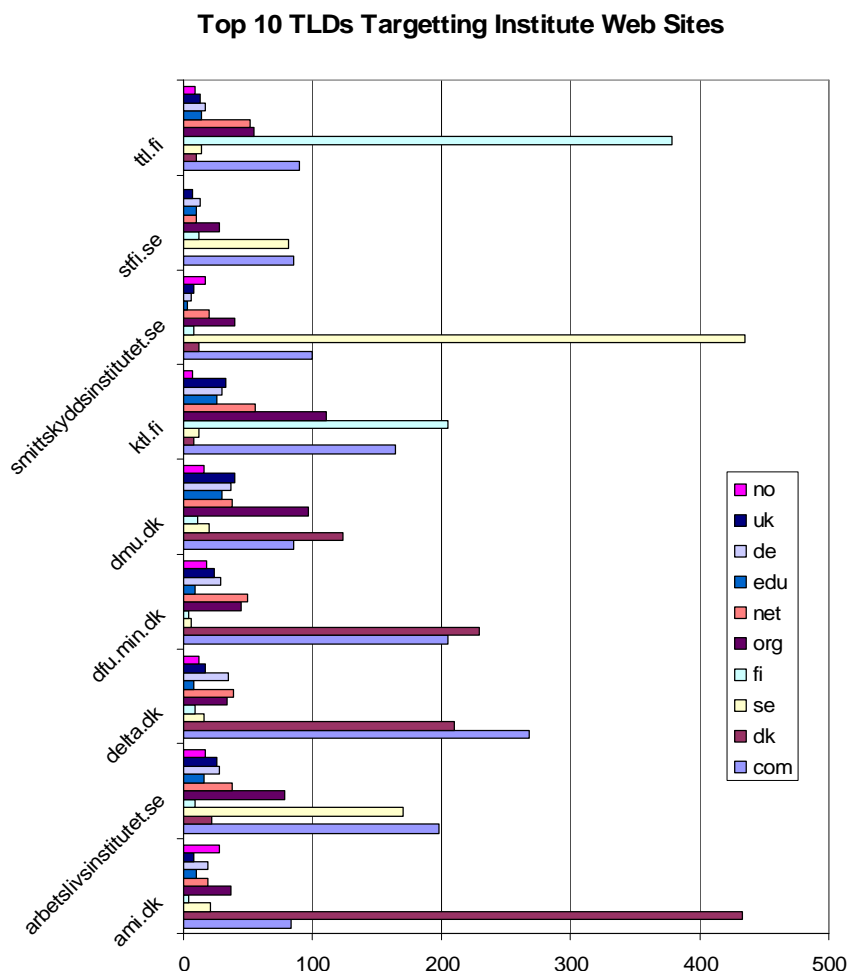


Figure 2. The web spheres of influence of Nordic research institutes.

Appendix 1: The EUROLABS project 2002–2003

The project was organised as a consortium of project groups led by the PREST group at Manchester University, under the project leader Luke Georghiou.

The **database** was designed to capture information available from documentary sources (which is subsequently validated). Key items covered include contact information, ownership, financial income, relationship with other organisations, location and structure, functions, sectors addressed, scientific and technological capability and personnel data. 769 centres were included in the database.

Data collection comprised three stages. The first stage made use of existing data sources. Published sources of data were gathered directly from the research centres. Secondly an Internet based ‘pre-survey’ was used to confirm and collect further information on the identified research centres. In the final stage the data was printed and mailed to the research centre for validation. A response rate of 66% was obtained. Data were entered in an Access database.

The following **definition** was used for extended Nordic version of the database that now is developed at the Swedish Research Council: “Research institutes are organizations outside of 1) the university sector, and 2) outside of firms producing other goods and services. The organizations should mainly perform R&D (>50 % of total budget) and have more than 4 active researchers.”

The 769 centres in the EUROLABS database account for over 25 billion EURO per annum (599 cases) in their budgets and employ over 100,000 scientists (557 cases). The greatest number of centres (237) employ between 10 and 49 scientists comprising 6% of the total number of scientists. The greatest number of scientists work either in large organisations (45,241 in 18 organisations comprising 43% of the total) or in centres with 100 to 499 scientists (33,785 in 151 centres).

References

- Bar-Ilan, J. (2001). Data collection methods on the Web for informetric purposes - A review and analysis. *Scientometrics*, 50(1), 7-32.
- Björneborn, L., & Ingwersen, P. (2001). Perspectives of Webometrics. *Scientometrics*, 50(1), 65-82.
- Brin, S., & Page, L. (1998). The anatomy of a large scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30(1-7), 107-117.
- Davenport, E., & Cronin, B. (2000). The Citation network as a Prototype for Representing Trust In Virtual Environments. In B. Cronin & H. B. Atkins (Eds.), *The Web of Knowledge: a Festschrift in Honor of Eugene Garfield*. (pp. 517-534). Metford, NJ: Information Today Inc. ASIS Monograph Series.
- Ingwersen, P. (1998). The calculation of Web Impact Factors. *Journal of Documentation*, 54(2), 236-243.
- Thelwall, M. (2004). Weak benchmarking indicators for formative and semi-evaluative assessment of research. *Research Evaluation*, 13(1), 63-68.
- Wilkinson, D., Harries, G., Thelwall, M., & Price, E. (2003). Motivations for academic Web site interlinking: Evidence for the Web as a novel source of information on informal scholarly communication. *Journal of Information Science*, 29(1), 49-56.