

What can University-to-Government Web links Reveal about University – Government Collaboration?

David Stuart and Mike Thelwall

dp_stuart@hotmail.com, m.thelwall@wlv.ac.uk

School of Computing and Information Technology, University of Wolverhampton,
Wulfruna Street, Wolverhampton WV1 1SB, UK.

Abstract

Recent link analysis research has found strong inter-university academic linking patterns. This study looks beyond the academic community, focusing on university-government linking practices. It finds a stronger correlation between a university's research productivity and the total number of outlinks than with the total number of university-government links. A classification exercise shows that university-to-government links predominately identify information resources rather than reflecting a research relationship.

Introduction

This study analyses university-to-government web links, investigating whether the number of university-to-government web links can be used as an indicator of a university's research productivity; and what the reasons for the university-to-government link placement tell us about university-government relationships. Over recent years, there have been a number of acknowledged changes in the way contemporary research is carried out, with greater emphasis being placed on collaborative and government-driven, commercially-oriented research (Gibbons et al., 1994; Etzkowitz & Leydesdorff, 2000). In this context, webometric indicators may be useful to show collaborations that may not be reflected in traditional bibliometrics (Thelwall, 2004a).

It is reasonable to question whether university-industry-government relationships are manifested on the World Wide Web, and if they are quantifiable in terms of the number of links from one organisation to another. Studies have already found web links to reflect real world phenomena (Thelwall, 2002), and Leydesdorff and Curran (2000) have shown that university-industry-government relations can be investigated with the use of advanced search engines. This research focuses on the relationship between United Kingdom universities and government departments, as manifested by the web links from university web sites to government web sites. University-to-government links are obtainable from a manageable number of easily identifiable university web sites, with links to government sites easily identifiable through their '.gov.uk' second-level domain name.

Investigating universities in the United Kingdom also allows for the number of links to be compared to an acknowledged indicator of research productivity, the Research Assessment Exercise (RAE, 2005).

Methodology

The investigation was carried out in three parts: first, establishing whether there was a correlation between the number of university-to-government web links and a university's research success; second, establishing whether this correlation was stronger than a correlation between the total outlinks from a web site and a university's research success; and finally, a classification of a sample of the university-to-government web links to establish the reasons for link placement. This is broadly in line with a recommended link analysis methodology (Thelwall, 2004b).

The data set was taken from a publicly available database containing the link structure of 125 UK universities (<http://cybermetrics.wlv.ac.uk/database/>). This database is based on a web crawler designed for the collection of data for the academic community (Thelwall, 2003a).

Establishing a University's Research Quality

A figure to express the research quality of a university was calculated based on the results of the 2001 Research Assessment Exercise (i.e., the most recent one). These scores were taken to be on a linear

scale, which is the standard interpretation (Education Guardian, 2001) and has been used previously in academic research (Thelwall 2002; 2004b). These scores were then multiplied by the number of researchers in the department, totalling the products for all the departments in a university. This total was then divided by the number of full time equivalent academic staff (Education Guardian, 2001; Thelwall & Harries, 2004). Such normalisation is essential because UK universities have greatly differing sizes: correlations for data without normalisation could be explained by size differences.

A program calculated the number of university-to-government web links for each university. This figure was normalised by dividing by the full time equivalent number of staff, and then compared to the university's research quality figure using Spearman's rank correlation test. This procedure was repeated for the total number of outlinks for each university.

Classification of reasons for hyperlinks

A classification of the hyperlinks was carried out using content analysis techniques, (Krippendorff, 2004). Following Wilkinson et al. (2003), a random selection of ten web links was extracted from each of the 125 universities' lists of web links to create a single file of random university-to-government web links that would not be heavily biased towards larger universities. For those universities where there were not ten links to government web sites, as many links as were available were added to the file. This file was then randomly sorted, and the first 400 were used as part of the classification exercise.

Although classifications of university hyperlinks have been carried out previously (Thelwall, 2003b; Wilkinson et al., 2003; Bar-Ilan, 2004), the classification schemes reflect the specific research questions addressed, and as such are inappropriate for the classification of university-to-government web links. The classification initially used three broad headings, links were attributed as representing an arbitrary relationship; an existing research based relationship; or a desire for there to be a relationship. Further breakdowns in the classification system were integrated in an inductive system, with finer grained classification being integrated as each of the pages were visited. When a link covered more than one of the categories, it was classified under the section that represented the highest level of research relationship that the page included.

Classification of links was carried out by examining the source web page to establish the link creation context. If necessary the context of the page was determined by looking at whereabouts the web page fitted into the university's web site. For any pages where it was not possible to classify the link by looking at the current page due to either the web page or the link no longer being present, a cached copy of the page was retrieved from the Internet Archive's Wayback machine (Internet Archive, 2005).

To reduce classifier bias more than one person should carry out any classification exercise, but a single classifier is acceptable for exploratory research (Thelwall, 2004b).

Results

Spearman's rank correlation coefficient for the correlation between the number of university-to-government web links per full time equivalent member of staff and the research productivity per member of staff gave a value of .168 ($n = 125$); which is not statistically significant. Spearman's rank correlation coefficient for the total number of outlinks was .738, which is statistically significant at the 1% level. Figures 1 and 2 express the correlations in the form of a scatter graph. Whilst the correlation is clearly visible in Figure 2, none is apparent in Figure 1.

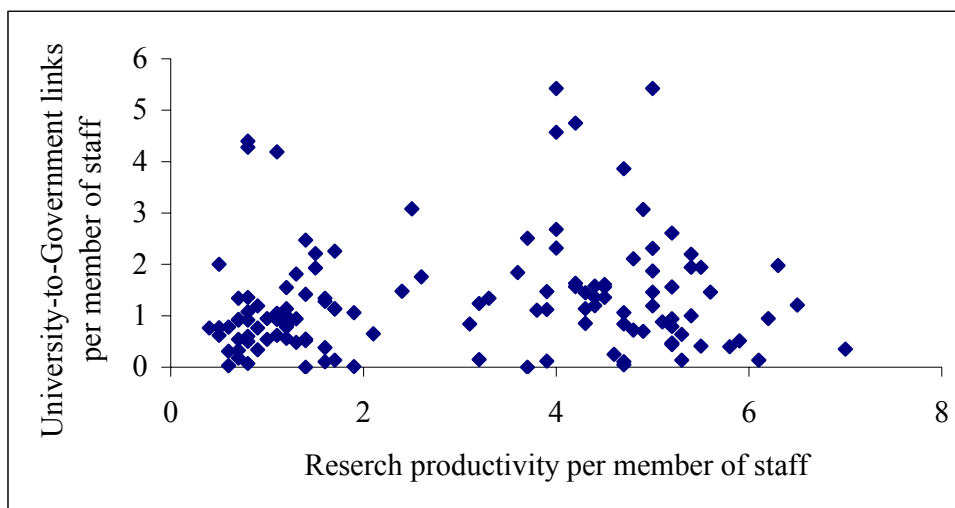


Figure 1: Scatter graph comparing the number of University-to-Government links with research productivity per member of staff

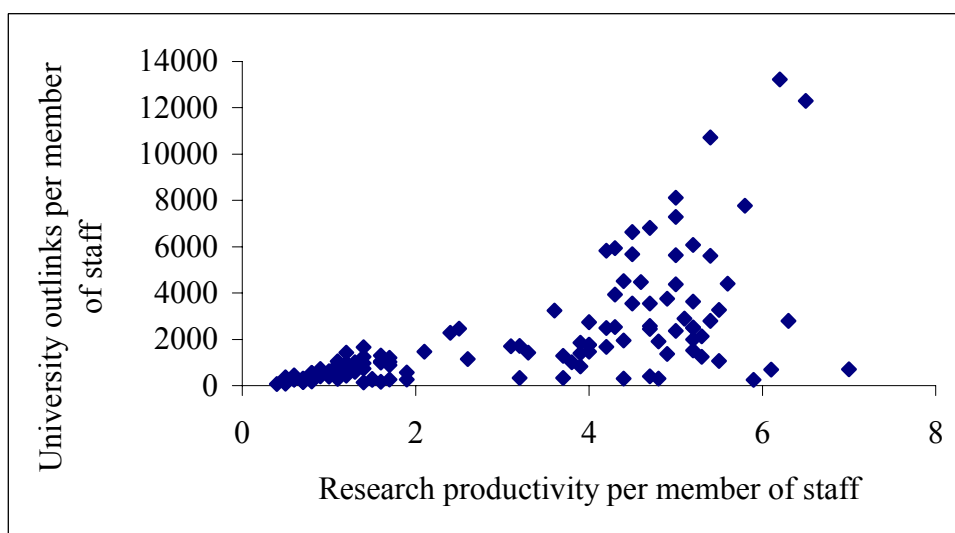


Figure 2: Scatter graph comparing total university outlinks and research productivity per member of staff

Table 1 shows the final classification scheme. For those links that could not be found either on the internet or on the Internet Archive, arguably the transitory nature of the link means that the relationship is unlikely to be permanent.

Table 1. Classification scheme for web links including the number of links in each category

Relationship inferred from web links	Number of links
1. Expressing a research relationship	
Previous work	2
A client	8
Funding / Sponsorship	12
Collaboration	13
Undisclosed	7
2. Expressing desire for relationship:	
Careers	16
Funding information	9
Other	2
3. Expressing an arbitrary relationship	
Academic resource	190
References	37
Other information	84
Link placed in error	2
4. Unable to determine relationship	18

Discussion

The results show a strong correlation between a university's research productivity and its total outlinks. Interestingly, the university's number of university-to-government links does not show a better correlation, and is not even statistically significant. There are a number of possible reasons for this:

- A greater percentage of university-to-non-government links reflect a research relationship than university-to-government links.
- University websites don't generally indicate the governmental origin of their research funding.
- The university-government relationship is not as important as other relationships to a university's research productivity.
- There were too few links for a pattern to emerge.

In fact, most UK government funding is indirect, via funding councils. The funding councils typically have .ac.uk academic domain names. The government's role is the provision of the funding, and the overall direction in which the funding should be allocated. Hence the lack of university-government links may reflect that the government does not actually exist as a serious part of the research funding process: the key players are the academics who decide upon the allocation of funds (Whitley, 2000). In other words, the university-government relationship is actually a university-university relationship in many ways, particularly from the perspective of individual researchers.

Conclusion

This research shows that whilst there is a correlation between a university's research productivity and the total number of outlinks, there is not a strong correlation between a university's research productivity and the number of university-to-government links. A classification of the links suggests that the reason for this is that the links are placed predominantly to highlight information resources and not real collaborative efforts.

The obscuring of the university-government connection has also been discussed as a possible cause of the results, with the government hidden behind research councils, which themselves rely heavily upon academics for their decisions.

It is likely that between each pair of top-level domains there are an identifiable set of reasons for the placing of web links, and these reasons differ in proportion for each. Further research is needed to identify these differences so that webometrics can be used to provide indicators to more specific questions.

References

- Bar-Ilan, J. (2004) A microscopic link analysis of academic institutions within a country - the case of Israel. *Scientometrics*, 59(3), 391-403.
- Education Guardian (2001) *About the tables*. Retrieved January 5, 2005, from <http://education.guardian.co.uk/secondaryschools/guide/story/0,11228,602663,00.html>.
- Etzkowitz, H., & Leydesdorff, L. (2000) The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university-industry-government relations. *Research Policy*, 29, 109-123.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994) *The new production of knowledge*. London: Sage.
- Internet Archive (2005) [*Homepage*]. Retrieved January 10, 2005, from <http://www.archive.org>.
- Krippendorff, K. (2004) *Content Analysis, an introduction to its methodology*. 2nd ed., London: Sage.
- Leydesdorff, L. & Curran, M. (2000) Mapping University-Industry-Government relations on the Internet: The construction of Indicators for a Knowledge-Based Economy. *CYBERmetrics*, 4. Retrieved January 7, 2005, from <http://cybermetrics.cindoc.csic.es/cybermetrics/articles/v4i1p2.html>.
- RAE (2005) *RAE 2008*. Retrieved January 7, 2005, from <http://www.rae.ac.uk/default.htm>.
- Thelwall, M. (2002) Evidence for the existence of geographic trends in university Web site interlinking. *Journal of Documentation*, 58(5), 563-574.
- Thelwall, M. (2003a) A free database of University Web Links: Data collection issue. *CYBERmetrics*, 6/7. Retrieved January 7th 2005 from <http://cybermetrics.cindoc.csic.es/cybermetrics/articles/v6i1p2.html>.
- Thelwall, M. (2003b, April) What is this link doing here? Beginning a fine-grained process of identifying reasons for academic hyperlink creation. *Information Research*, 8(3). Retrieved January 7, 2005, from <http://informationr.net/ir/8-3/paper151.html>.
- Thelwall M. (2004a) Non-robust indicator theory: Weak benchmarking indicators for formative and semi-evaluative assessment of research. *Research Evaluation*, 13(1), 63-68.
- Thelwall, M. (2004b). *Link Analysis: An Information Science Approach*. San Diego, CA: Academic Press.
- Thelwall, M. & Harries, G. (2004) Do the web sites of higher rated scholars have significantly more online impact? *Journal of the American Society for Information Science and Technology*, 55(2), 149-159.
- Whitley, R. (2000). *The Intellectual and Social Organization of the Sciences* (2 ed.). Oxford: Oxford University Press.
- Wilkinson, D., Harries, G., Thelwall, M. and Price, L. (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*, 23(1), 49-56.