JIS

# Webometric analysis of departments of librarianship and information science: a follow-up study

**Mónica Arakaki**

*University of Sheffield, United Kingdom*

**Peter Willett**[1]

*University of Sheffield, United Kingdom*

**Abstract**

**This paper reports an analysis of the websites of UK departments of library and information science. Inlink counts of these websites revealed no statistically significant correlation with the quality of the research carried out by these departments, as quantified using departmental grades in the 2001 Research Assessment Exercise and citations in Google Scholar to publications submitted for that Exercise. Reasons for this lack of correlation include: difficulties in disambiguating departmental websites from larger institutional structures; the relatively small amount of research-related material in departmental websites; and limitations in the ways that current Web search engines process linkages to URLs. It is concluded that departmental-level webometric analyses do not at present provide an appropriate technique for evaluating academic research quality, and, more generally, that standards are needed for the formatting of URLs if inlinks are to become firmly established as a tool for website analysis.**

Keywords: Citation analysis; Research Assessment Exercise; Webometrics

## 1. Introduction

Citation analysis has been used for many years to probe a range of phenomena in the information sciences, such as mapping the development of novel research areas, identifying the formal and informal links between researchers in a subject, and assessing research impact [1-4]. The advent of the Web soon led to the suggestion that "sitations" or "inlinks", i.e., links between Web sites, could be used in much the same way as conventional bibliographic citations [5,6], these early studies leading to the concept of webometrics [7-9].

Citation analyses have been carried out at multiple levels, with studies of the numbers of citations to individuals, research groups, departments, institutions and even to countries. Similar approaches were rapidly developed for the analysis of inlink data [10]. An early such study was that of Thomas and Willett [11], who

[1] Correspondence to Peter Willett, Department of Information Studies, University of Sheffield, 211 Portobello Street, Sheffield S1 4DP, United Kingdom, p.willett@sheffield.ac.uk.

investigated the websites of 14 library and information science (LIS) departments in the UK. This investigation, hereafter the Thomas study, included an attempt to correlate the number of inlinks to a department's website with that department's research quality, as evidenced by the grade that it achieved in the 1996 Research Assessment Exercise (RAE) [12]. The RAE is a regular review of the research activities of every department in every university in the UK, and it provides one of the principal determinants for the allocation of governmental funding to higher education institutions. The RAE is based on extensive peer review and results in each department achieving a numeric grade, the top grade representing "Quality that equates to attainable levels of international excellence in more than half of the research activity submitted and attainable levels of national excellence in the remainder" [13]. RAE grades have been found to correlate well with conventional citation counts, not just for LIS [14-16] but also for many other subjects, e.g., anatomy, archaeology and genetics [17]; and consideration is thus now being given to reducing the costs of the RAE by using performance metrics based on bibliometric data as a complement, or even an alternative, to the current procedures based on peer review [18, 19].

The Thomas study was unable to identify any statistically significant relationship between inlink counts and research excellence (as quantified by RAE grades), but subsequent studies have been more successful in demonstrating significant relationships at both institutional and departmental levels [20-24]. The aim of the work reported here was to update the Thomas study, with two specific objectives: to use more recent, and more detailed, RAE data than had been employed in that study; and to determine whether a significant correlation between inlink counts and research quality could now be identified. Full details of the experiments that were carried out here are presented by Arakaki [25].

## 2. Methods

The focus of the study was the departments that submitted themselves to Unit of Assessment (UoA) 61 (Library and Information Management) for assessment in the 2001 Research Assessment Exercise (RAE2001). The Thomas study (published in 2000 but carried out in 1999) involved the websites of 14 LIS departments; since this study, however, many of the departments examined have undergone radical changes in structure, merging with other departments or being subsumed within a larger organisational grouping such as a school or faculty (a process mirroring the situation in the USA [26]). At the same time, there are departments that submitted to RAE2001 UoA 61 that were not considered in the Thomas study, these often focusing their teaching and research on various aspects of the closely related topic of information systems. Of the 23 institutions submitting to UoA 61, Queen Margaret University College and the University of Central England (now Birmingham City University) no longer offer LIS courses and we hence discuss only the remaining 21 organisational units, as listed in Table 1. In what follows, we shall generally refer to these units as "departments", despite the fact (as discussed below) that some of them are actually schools or faculties.

***Table 1 near here***

### *2.1. RAE grade and citation data*

The Thomas study used RAE grades as a measure of the quality of a department's research, as have many other UK studies of research impact. A further RAE (RAE2001) has taken place since the Thomas study but the grades from this are now dated and reflect departmental research achievements in the period January 1994 – December 2000 (for subjects in the arts and humanities including LIS; for other subjects the period covered by RAE2001 is January 1996 – December 2000). Thus, while we have used the RAE2001 grades as a measure of research quality, we have also used a different measure that, we suggest, may better reflect the recent and current impact of a department's research. An important component of each department's RAE2001 submission was a list of (normally) four research outputs (e.g., journal articles, reports, books, etc.) for each member of staff included in the submission, these outputs being chosen to best reflect the research carried out by that individual in the period 1994-2000. The selected outputs for each department are available on the RAE2001 website (at http://www.hero.ac.uk/rae/), and we have used each output in turn for a citation search covering the period 2001 to mid-2007, thus giving the total number of citations to each department's chosen outputs since RAE2001. The outputs themselves, like the RAE grades, hence correspond to the pre-2001 period; however, the citations to those

outputs reflect the impact post-2001 and hence, we suggest, provide a more current measure of research quality than do the RAE2001 grades.

Citation studies have traditionally used the Thomson Corp. Web of Knowledge database since this was, for many years, the only available source of citation data. The last few years have seen the advent of two further data sources that contain large amounts of citation data - Google Scholar and Scopus (see, e.g., [27,28]) – and we have chosen to use Google Scholar for the analyses reported here. The principal reason for this choice is that Web of Knowledge and Scopus focus on the journal literature, whereas a fair number of the RAE2001 LIS outputs came from conferences or other non-journal sources. Google Scholar is based on the Web (or at least that part of it that is perceived by Google's indexing software), and it thus provides better citation coverage of non-journal materials than does the Web of Knowledge. There is now substantial evidence that Google Scholar is a more appropriate data source for disciplines that make extensive use of the non-journal literature [29, 30], and it has been benchmarked for the purposes of research evaluation against the university grades resulting from the Performance-Based Research Fund 2003 assessment exercise that was carried out in New Zealand [31].

The 21 departments in Table 1 returned a total of 301 members of staff, these submitting a total of 1161 research outputs. The citations for each output were checked in Google Scholar, with the searches being carried out between 11th June and 13th July, 2007. The citations for each output were inspected carefully, since it soon became clear that there were numerous duplicate entries (typically caused by variations in the names of authors or journals or in the titles of the outputs) and that it was sometimes necessary to use both American and British spellings to retrieve all of the citations. The limitations of Google Scholar for large-scale bibliometric analyses are described in detail by Meho and Yang [28].

## *2.2. Inlink data*

The numbers of inlinks to the websites of each of the departments were obtained by running LexiURL Searcher, a software program developed by the Statistical Cybermetrics Research Group in the School of Computing and IT at the University of Wolverhampton. This tool, among others, is freely downloadable from the research group's website at http://cybermetrics.wlv.ac.uk. LexiURL Searcher acts as an intermediary by sending user-defined queries through the Applications Programming Interface (API) of an operational Web search engine (Google and Yahoo! when this work was carried out, but now also Live Search) and then returns the search results to the user for inspection and/or subsequent analysis. Of the two search engines that were available, Google provides a link operator that returns the number of inlinks to a specified site: however, it only displays a subset of these sites, apparently so as to prevent webmasters from manipulating their sites' rankings [32]. Accordingly Yahoo! was chosen as the research tool to be used in combination with LexiURL Searcher: LexiURL Searcher was preferred to Yahoo! Site Explorer for this part of the work since the former better supports complex queries and generates an output file of the inlinking sites for subsequent processing.

An example of a Yahoo! query is:

(link:http://www.slais.ucl.ac.uk OR link:http://www.ucl.ac.uk/slais) -site:ucl.ac.uk

This query retrieves the pages with links pointing to the website of the School of Library, Archive and Information Studies at University College London (UCL). The initial, bracketed part of the query reflects the fact that the website can be accessed through two different URLs, while the final part preceded by the minus operator is used to remove self-links (i.e., those generated from any page under the ucl.ac.uk domain). The availability of two different URLs - one in which the domain name corresponds to the department itself (http://www.slais.ucl.ac.uk) and the other in which the domain name corresponds to the parent organisation (http://www.ucl.ac.uk/slais) – is of considerable significance, as discussed further below. In like manner, some departments used alternative domain names, e.g., bton.ac.uk and brighton.ac.uk: both alternatives were again included in the search in such cases. A complete Yahoo! query is stored as a text file that acts as the input to LexiURL Searcher: the program sends the queries to Yahoo!, receives the results and then generates a text file containing the sites that linked to the SLAIS website.

### *2.3. Content analysis*

The two sets of data - Google Scholar citation counts and LexiURL Searcher inlink counts - were processed using SPSS, with the Spearman rank correlation coefficient being used to quantify the extent of the relationships between the variables that were considered. Following the correlation study, samples of the websites that linked to LIS departments' websites were examined to identify the natures of the information that was stored and of the linkages that had been made. Specifically, the pages that linked to the De Montfort, London, Sheffield and Strathclyde sites were downloaded from Yahoo! Site Explorer, and 30 pages chosen at random, except in the case of De Montfort which only received 19 inlinks[2]. If it was possible to reach the departmental online information through more than one URL, the sample included a proportional combination of pages pointing to each of them.

## 3. Results

The results of the citation and inlink analyses are listed in Table 2, together with the RAE2001 grades. The table is ordered in decreasing order of the total number of citations in Google Scholar to the research outputs submitted by each of the 21 departments (the third column in the table); slightly different rankings are obtained if the normalised citation counts (i.e., the mean values when averaged over the numbers of staff or over the numbers of outputs) are used. As would be expected, there is a well-marked skewed distribution, with a few of the institutions attracting the great bulk of the citations: thus the top four departments in Table 2 received 6933 (61.8%) of the citations despite having only 96 (31.9%) of the staff submitted to the RAE. Indeed, the skewness is also marked within a department, with five cases (Aberystwyth, Glasgow, Staffordshire, University College London, and West of England) where one individual's citations represented at least 50% of the total citations for that department.

As noted in the Introduction, several previous studies have demonstrated strong correlations between citation counts and RAE scores, and this is again the case here. The values for the Spearman rank correlation coefficient for the correlation between RAE grade and total citations, mean citations per staff-member and mean citations per output are 0.69, 0.69 and 0.70, respectively; all of these results are statistically significant at the 0.01 level of statistical significance. For comparison, two previous studies (using slightly different methodologies) of UK LIS departments considered correlations with the RAE1992 grades: Oppenheim's study of 13 departments reported values of 0.81 and 0.82 for the Spearman correlations between RAE grades and total citations and citations per staff member, respectively [15]; Seng and Willett's study of seven departments reported values of 0.95 and 0.95 for these two Spearman correlations [14]. The fact that the citations for the sets of nominated outputs are well correlated with the RAE2001 grades suggests that those departments that performed well in RAE2001 have continued to exert a high level of impact since 2001. It is hence not unreasonable to surmise that the same departments will also perform well when the results of the next RAE are announced in late-2008.

***Table 2 near here***

The final column of Table 2 lists the inlink counts for each of the 21 departments. It is clear that the relationship with citations is much less well marked than in the case of the RAE grades: indeed, the value of the Spearman coefficient for the correlation between citations and inlinks is as low as 0.05 (with the value for the correlation between RAE grades and inlinks being 0.11). Inspection of Table 1 will reveal that the departments differ significantly in form and it was thought that this might have affected the inlink counts that were obtained. Specifically, the departments can be grouped into three broad classes depending on where LIS teaching and research is carried out: by a distinct unit within a school or faculty (e.g., the Department of Information Science is part of the Faculty of Science at Loughborough University); by a distinct department without any higher-level of organisation (the only example identified was the Department of Information Studies at the University of Wales, Aberystwyth); or by a faculty or school, without any subordinate level of organisation (e.g., the School of Computing, Mathematical and Information Sciences at the University of Brighton). Inspection of Table 1 shows

[2] De Montfort had received 20 inlinks according to Table 2. However, when conducting the content analysis a few weeks after gathering the initial data, this count was found to be 19 instead, i.e., there was at least some variation in the inlink counts during the period of the study. This has been ignored in the discussion.

that several of the last group are multi-component faculties covering a range of areas of study (e.g., Brighton mentioned previously), whereas two of them (De Montfort University and the University of Paisley) have faculties that involve just computing in their name. We have hence divided the various organisational units into two classes: "faculty-like" units that are in multi-subject faculties or schools (Leeds Metropolitan, Northumbria at Newcastle, Salford, Staffordshire, West of England) and "department-like" units that are based in specific named departments or single-subject faculties (the rest). Spearman coefficients were calculated using just the members of each of the two groups, with the resulting correlations between citations and inlinks being 0.14 (faculty-like) and 0.07 (department-like), i.e., only marginally larger than for the entire set of 21 departments.

None of the three Spearman values computed here is significant at the 0.05 level of statistical significance, and we hence conclude that there is no significant relationship between a department's research quality (as denoted by citations to its RAE-submitted publications) and its web presence (as denoted by inlinks to its website).

## 4. Discussion

The lack of a relationship between research impact and inlink counts was noted in the Thomas study. However, this was carried out at a time (1999) when many of the departmental websites studied were still at an early stage of development, and later studies by the Thelwall group identified significant relationships at both the institutional and the departmental levels [20-24]. Our results are thus very different from what had been expected when we started this study, and we now discuss three possible reasons for the observed behaviour.

We believe that the principal reason for our unexpected findings relates to the structure of URLs. Specifically, the inlink counts for a website are crucially dependent on the way that a department organises its website and on the interaction of this organisation with the current generation of search engines. The link operator that was used in our searches returns the number of inlinks to a specific ***page***, not to an entire ***site***; in other words, the link operator does not take into account pages located hierarchically below the specified URL. If the number of links to all pages within a site is required, the linkdomain operator should be used instead. However, this operator only works with URLs specified as site home pages (such as http://www.cis.strath.ac.uk or http://www.slais.ucl.ac.uk) and does not work with URLs that contain path names (such as http://www.strath.ac.uk/cis or http://www.ucl.ac.uk/slais). The latter situation was the case for most of websites listed in Table 1. Consequently, the figures given in Table 2 are inlinks to the departmental home pages, a situation that is clearly far from ideal, since much of a department's web impact might be due to subordinated pages lower in the hierarchy. To take an extreme example, there are just 516 inlinks to the homepage of the Department of Computer and Information Sciences at Strathclyde, whereas there are no less than 3865 inlinks to the entire site (comprising 20475 pages). It must be emphasised that this failure to include inlinks to subordinate pages is not a limitation of the LexiURL Searcher software: it is, instead, simply the way that current search engines process a search query. A related problem is that there may simply be more links to pages elsewhere in an academic hierarchy: for example, the department at Brunel received only a single inlink whereas there were 95 inlinks to its parent, the School of Information Systems, Computing and Mathematics, i.e., although there is a distinct department, it is the parent school that has attracted the inlinks..

Thus far, we have discussed this problem in purely technical terms, and there is no reason why changes in the way that search engines operate could not obviate this limitation of current systems. For example, Google has recently announced that it will enable inlink-based searching of entire sites [32], but the facility will only be made available to registered webmasters [33]. This change represents a backwards step as this facility was available to all in the AltaVista software that was used for most early webometrics studies (including those by the Thelwall group that have been cited previously): as Bar-Ilan has noted, the priorities of commercial search engine companies may not reflect those of academic researchers [34]. However, even allowing for future changes in search engine technology, there is a more general problem here, *viz* the lack of standards for the structuring of URLs. Bibliographic control, in all its manifestations, has been one of the key driving forces in the development of librarianship and then of information science, but it was only with the establishment of MARC and AACR that worldwide standards became established for the categorisation of bibliographic entities. Comparable efforts are now being made for the categorisation of web entities by the establishment of metadata standards. However, these efforts, as with conventional book cataloguing, have focused on the categorisation of individual items rather than

on the very large number of items that may comprise an individual website, despite the fact that inlinks to a website are a typical example of metadata, i.e., data about data. Just as there is no standard way of describing collections of bibliographic items, i.e., libraries, so there is no standard way of describing collections of digital items, i.e., websites, despite the huge and growing number of them that exist and despite the increasing interest in analysing them by purely automatic means. The present study has highlighted a specific consequence of this lack of standardisation, but the implications are more widespread. Higher education institutions will be keen to ensure that publications by their staff receive the maximum amount of visibility, especially if linkage-based tools were to be considered for use in future RAE-like analyses. The managers of these institutions would probably be surprised were they to realise that while the Department of Information Studies at the University of Sheffield and the University of Sheffield's Department of Information Studies, for example, are perceived as being synonymous in the real world, websites corresponding to these two modes of organization might be perceived very differently in the digital world. This difference clearly has implications for the institutional repositories that are increasingly being used to store academics' publications: it seems not unreasonable to expect the development of corporate strategies to maximise the visibility of such repositories and the documents therein, in much the same way as commercial organisations currently seek to maximise their rankings in search-engine outputs.

There are two other possible reasons for the lack of any significant relationship between research impact and inlink count.

First, we have noted previously that it may be difficult to disaggregate the pages representing a department from a larger organisational unit, and that a department's homepage might have alternative URLs. Related to the latter is the fact that a department may also have other sites associated with it in a manner that would be unlikely to be identified from the homepage by any search engine. For example, much of the research in information retrieval and in librarianship carried out in the authors' department in Sheffield is described at http://ir.shef.ac.uk/ and http://cplis.shef.ac.uk/, respectively, locations that are not obviously linked to the official departmental homepage at http://www.shef.ac.uk/is. In fact, the situation in Sheffield is still further complicated by an alternative site at http://dis.shef.ac.uk. Yahoo! Site Explorer identifies 134 inlinks to the departmental homepage but no less than 1751 inlinks to this alternative homepage, which is for a developmental server that has significant numbers of subordinate pages that are highly sited (e.g., the Information Literacy Place at http://dis.shef.ac.uk/literacy has 199 inlinks, 65 more than the official departmental homepage itself). We have not included these inlinks in the counts in Table 2 since they would bias the Sheffield count as compared to the other departments, which may well have similar associated sites of which we were unaware. Such occurrences, which in this case at least reflect the organisational structure of the servers providing access to the data, again highlight the need for standard naming procedures that has been discussed previously.

Second, we must consider the nature of the linkages that make up the inlink counts. Four departments were selected for this purpose: De Montfort, UCL, Sheffield and Strathclyde, as described in the Methods section. Considering the inlinks to De Montfort, only three of the linking pages (16%) represented linkages based on academic content: an announcement about a workshop on computing (from the University of Leicester), news of satellite software developed at De Montfort (from a German TV provider), and a blog about biomedical information. All of the other 16 linkages to the De Montfort site were from companies that maintained mirrored data of the Open Directory Project. The inlinks to UCL were very different, with no less than 87% of them seemingly academic in nature. Examples of linking organisations included the Society of Qualified Archivists, the Scientific Archivists Group, and the Association for Literary and Linguistic Computing, the National Archives in the UK and Canada, and a range of sites related to various aspects of humanities computing. Thus the great majority of the UCL inlinks involved topics where the department has a well-established research reputation. The percentage of academically-related inlinks was still greater (93%) for Sheffield. However, whilst there were some pages related to the department's research activities (e.g., blogs about the invisible web and research methodologies) there were many others that did not come into that category (e.g., online CVs and university library directories). Similar comments apply to Strathclyde: of the 70% of academic inlinks, the largest contributions came from pages regarding a KDE (a graphical desktop environment for Linux and UNIX workstations) convention that was hosted by Strathclyde in 2007 and from blogs on general computer science topics. We have used only four of the departments for the content analysis but the observed lack of fit between content and inlinks was noted in the early days of webometrics [11, 35] and things have clearly changed little since then. It is still an open question as to why this is so, despite repeated requests for studies of the reasons for hyperlinking [36, 37], whereas there have been many such studies of the reasons for citing (as reviewed recently by Bornmann and Daniel [38]).

To summarise: current data collection tools may not count what is actually required (i.e., inlinks to sites rather than to pages); departmental websites cannot always be identified sufficiently precisely for the purposes of quantitative analysis; and inlinking is often carried out for purposes unrelated to academic research. Taken together, these three factors mean that the inlink counts for departmental websites are most unlikely to provide a robust measure of the quality of departmental research. A similar lack of correlation has been noted for European life-sciences research groups, rather than whole departments, in very recent work by Barjak and Thelwall [39]. A further problem, not studied here, is that there would appear to be substantial differences between the natures of departmental websites in different subjects and countries [40], which would further complicate the use of inlink data for multidisciplinary and multinational evaluations of research activity.

## 5. Conclusions

This paper has sought to update an earlier study of inlinks to the websites of LIS departments in the UK and of the relationship between inlinks and RAE data. We have been unable to identify any such relationship owing to the very incomplete, and potentially misleading, nature of the inlink data that can be collected at the departmental level (although this may be less of a problem when data is aggregated over an institution or country). We hence suggest that numbers of inlinks (as computed using currently available tools) do not provide an appropriate performance indicator for the quantification of academic research impact at the departmental level. We also suggest that at least some of the problems we have identified could be alleviated by the development of standards for the naming of websites.

**Acknowledgements**. We thank Prof. Mike Thelwall for provision of the LexiURL Searcher software and for advice on its use; and him, Prof. Nigel Ford and the referees for helpful comments on an earlier draft of this paper.

## 6. References

[1] E. Garfield, *Citation Indexing -- Its Theory and Application in Science, Technology, and Humanities* (ISI Press, Philadelphia, PA 1979).

[2] B. Cronin, *The Citation Process. The Role and Significance of Citations in Scientific Communication* (Taylor Graham, London, 1984).

[3] C.L. Borgman and J. Furner, Scholarly communication and bibliometrics, *Annual Review of Information Science and Technology* 36 (2002) 3-72.

[4] J. Nicolaisen, Citation analysis, *Annual Review of Information Science and Technology* 41 (2007) 609-641.

[5] R.R. Larson, Bibliometrics of the World Wide Web: an exploratory analysis of the intellectual structure of cyberspace, *Proceedings of the Annual Meeting of the American Society for Information Science* 33 (1996) 71-78.

[6] G. McKiernan, CitedSites(sm): Citation indexing of Web resources (1996) at URL http://www.public.iastate.edu/~CYBERSTACKS/Cited.htm [Accessed 19th February 2008]

[7] T.C. Almind and P. Ingwersen, Informetric analyses on the World Wide Web: methodological approaches to 'Webometrics', *Journal of Documentation* 53 (1997) 404-426.

[8] R. Rousseau, Sitations: an exploratory study, *Cybermetrics*, 1 (1997) at URL http://www.cindoc.csic.es/cybermetrics/articles/v1i1p1.html [Accessed 19th February 2008].

[9] P. Ingwersen, The calculation of Web impact factors, *Journal of Documentation* 54 (1998) 236-243.

[10] M. Thelwall, L. Vaughan and L. Björneborn, Webometrics, *Annual Review of Information Science and Technology* 39 (2005) 81-135.

[11] O. Thomas and P. Willett, Webometric analysis of departments of librarianship and information science *Journal of Information Science* 26 (2000) 421-428.

[12] HERO, *1996 Research Assessment Exercise* (1996) at URL http://www.hero.ac.uk/rae/rae96/ [Accessed 29th November 2007].

[13] HERO, *Research Assessment Exercise 2001. Guidance to Panel Chairs and Members: Criteria and Working Methods* (2001) at URL http://195.194.167.103/PanGuide/ [Accessed 19th February 2008].

[14] L.B. Seng and P. Willett, The citedness of publications by United Kingdom library schools, *Journal of Information Science* 21 (1995) 68-71.

[15] C. Oppenheim, The correlation between citation counts and the 1992 Research Assessment Exercise ratings for British Library and Information Science university departments, *Journal of Documentation,* 51 (1995) 18-27.

[16] A. Holmes and C. Oppenheimer, Use of citation analysis to predict the outcome of the 2001 Research Assessment Exercise for Unit of Assessment (UoA) 61: Library and Information Management, *Information Research* 6 (2) (2001) at URL http://informationr.net/ir/6-2/paper103.html [Accessed 29th November 2007].

[17] C. Oppenheim, The correlation between citation counts and the 1992 Research Assessment Exercise ratings for British research in genetics, anatomy and archaeology, *Journal of Documentation* 53 (1997) 477-487.

[18] Department for Children, School and Families, *Reform of Higher Education Research Assessment and Funding* (2006) at URL http://www.dfes.gov.uk/consultations/conResults.cfm?consultationId=1404 [Accessed 19th February 2008].

[19] British Academy, Response to the DfES Consultation on the Reform of Higher Education Research Assessment and Funding (2006) at URL http://www.britac.ac.uk/reports/rae-2006/response.html [Accessed 19th February 2008].

[20] A. Smith and M. Thelwall, Web impact factors for Australasian universities, *Scientometrics* 54 (2002) 363-380.

[21] X. Li, M. Thelwall, P. Musgrove and D. Wilkinson, The relationship between the WIFs or inlinks of computer science departments in UK and their RAE ratings or research productivities in 2001, *Scientometrics* 57 (2003) 239-255.

[22] M. Thelwall and G. Harries, The connection between the research of a university and counts of links to its web pages: an investigation based upon a classification of the relationships of pages to the research of the host university, *Journal of the American Society for Information Science and Technology* 54 (2003) 594-602.

[23] M. Thelwall and R. Tang, Disciplinary and linguistic considerations for academic Web linking: an exploratory hyperlink mediated study with Mainland China and Taiwan, *Scientometrics,* 58 (2003) 153-179.

[24] R. Tang and M. Thelwall, Patterns of national and international Web inlinks to US academic departments: an analysis of disciplinary variations, *Scientometrics* 60 (2004) 475-485.

[25] M. Arakaki, *How Visible are the UK Departments of Librarianship and Information Science? A Webometric Analysis*, MSc thesis, Department of Information Studies, University of Sheffield, 2007.

[26] C.R. Hildreth and M. Koenig, Organizational realignment of LIS programs in academia: from independent standalone units to incorporated programs, *Journal of Education for Library and Information Science*, 43 (2002) 126-133.

[27] P. Jacso, As we may search - comparison of major features of the Web of Science, Scopus, and Google Scholar citation-based and citation-enhanced databases, *Current Science* 89 (2005) 1537-1547.

[28] L. Meho and K. Yang, Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar, *Journal of the American Society for Information Science and Technology* 58 (2007) 2105-2125.

[29] K. Kousha and M. Thelwall, Google Scholar citations and Google Web/URL citations: a multi-discipline exploratory analysis, *Journal of the American Society for Information Science and Technology* 58 (2007) 1055-1065.

[30] M. Sanderson, Revisiting *h* measured on UK LIS and IR academics, *Journal of the American Society for Information Science and Technology* 59 (2008) 1184-1190.

[31] A.G. Smith, Benchmarking Google Scholar with the New Zealand PBRF research assessment exercise, *Scientometrics* 74 (2008) 309-316.

[32] D. Sullivan, Google releases new link reporting tools, at URL http://searchengineland.com/070205-165836.php [Accessed 22nd April 2008].

[33] Google, How can I see links to my site?, at URL http://www.google.com/support/webmasters/bin/answer.py?answer=55281&ctx=sibling [Accessed 22nd April 2008].

[34] J. Bar-Ilan, Data collection methods on the Web for informetric purposes: a review and analysis, *Scientometrics* 50 (2001) 7-32.

[35] M. Thelwall, The top 100 linked-to pages on UK university web sites: high in-link counts are not usually associated with quality content, *Journal of Information Science* 28 (2002) 483-491.

[36] H. Chu, Taxonomy of inlinked Web entities: What does it imply for webometric research?, *Library and Information Science Research 27* (2005) 8–27.

[37] P. Ingwersen and L. Björneborn, Methodological issues of webometric studies. In: H.F. Moed, W. Glänzel and U. Schmoch (eds.), *Handbook of Quantitative Science and Technology Research* (Kluwer, London, 2004) 339-369.

[38] L. Bornmann and H.-D. Daniel, What do citation counts measure? A review of studies on citing behavior, *Journal of Documentation* 64 (2008) 45-80.

[39] F. Barjak and M. Thelwall, A statistical analysis of the web presences of European life sciences research teams, *Journal of the American Society for Information Science and Technology* 59 (2008) 628-643.

[40] X. Li, M. Thelwall, D. Wilkinson and P. Musgrove, National and international university departmental Web site interlinking, *Scientometrics* 64 (2005) 151-185 (Part 1) and 187-208 (Part 2).

| Institution | School/Faculty | Department | URL |
|---|---|---|---|
| University of Brighton | School of Computing, Mathematical and Information Sciences | | http://www.brighton.ac.uk/cmis/<br>http://www.bton.ac.uk/cmis/ |
| Brunel University | School of Information Systems, Computing and Mathematics | Department of Information Systems and Computing | http://www.brunel.ac.uk/about/acad/siscm/disc |
| City University | School of Informatics | Department of Information Science | http://www.soi.city.ac.uk/organisation/is |
| De Montfort University | School of Computing | | http://www.dmu.ac.uk/faculties/cse/computing/ |
| Leeds Metropolita University | Innovation North Faculty of Information and Technology | | http://www.leedsmet.ac.uk/inn/<br>http://www.lmu.ac.uk/inn/ |
| Loughborough University | Faculty of Science | Department of Information Science | http://www.lboro.ac.uk/departments/dis/<br>http://www.lboro.ac.uk/departments/is/<br>http://www.lboro.ac.uk/departments/ls/ |
| Manchester Metropolitan University | Faculty of Humanities, Law & Social Science | Department of Information and Communications | http://www.hlss.mmu.ac.uk/infocomms<br>http://www.mmu.ac.uk/h-ss/dic/ |
| University of Northumbria at Newcastle | School of Computing, Engineering and Information Sciences | | http://northumbria.ac.uk/sd/academic/ceis/ |
| University of Salford | School of Computing, Science and Engineering | | http://www.cse.salford.ac.uk/ |
| University of Sheffield | Faculty of Pure Science | Department of Information Studies | http://www.sheffield.ac.uk/is<br>http://www.shef.ac.uk/is |
| South Bank University | Faculty of Business, Computing and Information Management | Department of Information Systems and Information Technology | http://www.lsbu.ac.uk/bcim/depts/isit/ |
| Staffordshire University | Faculty of Computing, Engineering and Technology | | http://www.fcet.staffs.ac.uk/<br>http://www.staffs.ac.uk/academic/comp_eng_tech/ |
| Thames Valley University | Faculty of Professional Studies | Department of Computing | http://www.tvu.ac.uk/academic_departments/Faculty_of_Professional_Studies/Computing.jsp |
| University College London | Faculty of Arts and Humanities | School of Library, Archive and Information Studies | http://www.slais.ucl.ac.uk/<br>http://www.ucl.ac.uk/slais |
| University of West of England, Bristol | Faculty of Computing, Engineering and Mathematical Sciences | | http://www.uwe.ac.uk/cems/ |
| University of Glasgow | Faculty of Arts | Humanities Advanced Technology and Information Institute | http://www.hatii.arts.gla.ac.uk |
| Napier University | Faculty of Engineering, Computing and Creative Industries | School of Computing | http://www.soc.napier.ac.uk/ |
| University of Paisley | School of Computing | | http://www.paisley.ac.uk/computing/ |
| Robert Gordon University | Aberdeen Business Schoo | Department of Information Management | http://www.rgu.ac.uk/abs/staff/page.cfm?pge=38794 |
| University of Strathclyde | Faculty of Science | Department of Computer and Information Sciences | http://www.cis.strath.ac.uk/<br>http://www.strath.ac.uk/cis/ |
| University of Wales, Aberystwyth | | Department of Information Studies | http://www.dis.aber.ac.uk<br>http://www.dil.aber.ac.uk |

*Table 1. Organisational units submitting to UoA 61*

| LIS Department | 2001 RAE grade | Total citations | Citations per staff member | Citations per output | Inlinks |
|---|---|---|---|---|---|
| Sheffield | 5* | 2305 | 109.76 | 28.81 | 134 |
| Brunel | 5 | 1873 | 58.53 | 15.10 | 1 |
| Salford | 5* | 1736 | 75.48 | 20.67 | 197 |
| City | 5 | 1019 | 50.95 | 13.06 | 124 |
| Loughborough | 5 | 537 | 28.26 | 7.26 | 301 |
| Strathclyde | 4 | 516 | 32.25 | 8.60 | 516 |
| Northumbria | 3b | 449 | 24.94 | 6.24 | 212 |
| Brighton | 3b | 423 | 23.50 | 6.41 | 191 |
| Wales | 3a | 316 | 18.59 | 4.79 | 198 |
| Staffordshire | 3a | 284 | 40.57 | 10.14 | 210 |
| Manchester | 4 | 269 | 17.93 | 4.48 | 181 |
| De Montfort | 3a | 261 | 32.63 | 8.70 | 20 |
| London | 4 | 235 | 15.67 | 3.92 | 188 |
| South Bank | 3b | 225 | 10.23 | 2.62 | 3 |
| Leeds | 4 | 175 | 19.44 | 4.86 | 50 |
| Robert Gordon | 3b | 173 | 21.63 | 5.41 | 8 |
| Napier | 4 | 143 | 23.83 | 6.50 | 290 |
| West of England | 3b | 132 | 13.20 | 3.47 | 636 |
| Paisley | 3b | 102 | 12.75 | 3.19 | 3 |
| Glasgow | 3a | 42 | 6.00 | 1.68 | 379 |
| Thames Valley | 1 | 12 | 6.00 | 1.50 | 9 |

***Table 2**. Citations in Google Scholar to RAE2001 outputs submitted by 21 HEIs to UoA 61, and inlinks to the LIS websites for those departments*