Jointly published by Akadémiai Kiadó, Budapest and Kluwer Academic Publishers, Dordrecht

Scientometrics, Vol. 54, No. 3 (2002) 363–380

Web Impact Factors for Australasian universities

ALASTAIR SMITH,[1] MIKE THELWALL[2]

[1]*School of Communications and Information Management, Victoria University of Wellington (New Zealand)*
[2]*School of Computing and Information Technology, University of Wolverhampton (UK)*

Counts of links into the websites of Australasian universities were calculated from the output of a specially designed crawler that covered universities in the UK, Australia and New Zealand. These figures were compared to those from the commercial search engines AltaVista and AllTheWeb. Web Impact Factors (WIFs) for Australasian universities were then calculated by dividing link counts from the three countries by academic staff numbers at each target university. WIFs were compared with each other and also with a conventional measure of research output for Australia. It was discovered that the crawler-generated link counts were roughly proportional to those from AltaVista and AllTheWeb for Australia, albeit with some outliers in the data. WIFs correlated quite well with research output for Australia, but the relationship was not clear enough to be able to differentiate between the characters of the WIFs from the three different sources. However, a new measurement introduced, the normalised propensity to link, suggests that the New Zealand university web is more insular than that of Australia.

Introduction

The World Wide Web provides a fertile ground for the extension of the bibliometric techniques developed for the print environment. A number of studies have been published, establishing the concept of "webometrics" – the application of bibliometric techniques to the Web.[1-3] In particular, the Web Impact Factor (WIF) has been proposed[4] as a method of gauging the influence of web sites on each other, and of analysing links between institutions. This study investigates the use of WIFs to compare the research links between Australian, New Zealand and British universities and introduces a new related measure. Ingwersen proposed web Impact Factors by analogy with journal Impact Factors. The journal Impact Factor can be broadly defined as the ratio of a numerator: the citations made to a journal, to the denominator: the number of articles (or citable items) published in that journal.[5] In the Web Impact Factor, on the other hand, the numerator is the number of links made to a web site and the denominator is a measure of the size of the site. In Ingwersen's original proposal, the denominator was the number of pages at the site. There are problems with this approach though, and

0138–9130/2002/US $ 15.00

it has been proposed[6] that for a university the number of academic staff at the site is a better indicator of the overall size of the organisation. This is because the number of pages at a site is affected by factors unrelated to the size of the institution or its research output. For instance a document may be presented as one large web page, or broken up into smaller screen size pages. It is possible that other measures could be used for the size of the site, for instance the overall budget of the university, the number of equivalent full time students, etc. However number of academic staff is a reasonable measure of the capacity of an institution to carry out research. The original version of this paper presented at the ISSI conference[7] compared the two main different approaches and discussed the methodological issues involved, finding that the results of both were similar, but this update will use academic staff numbers alone and focus upon the outcome of the calculations instead.

The most convenient way of measuring links between institutions is to use the advanced search facilities of commercial search engines, several of which include link data in their databases. Indeed the Google search engine uses link data in its results ranking and it appears likely that others include this information in their proprietary algorithms. Several WIF studies (e.g. Ref. 8) have been carried out using the advanced search features of the AltaVista search engine. However commercial search engines are restricted in their coverage[9] and are optimised for searching rather than for webometric studies. To circumvent these problems a specialised crawler has been developed.[6] In this study, it will be used as the primary source of information, but its performance will be compared with two commercial search engines: AltaVista and AllTheWeb to provide an indication of the extent to which the results may be dependent on the data source.

Web link analysis of university sites is promising in general terms, but in practice web links are not entirely equivalent to citations in the scholarly literature. Whilst much of the web content of university sites is research related, there is significant content that is administrative, teaching, and recreational in nature. This is reflected in the links made to a site. For instance a significant proportion of the links made to the Victoria University of Wellington are to a popular database of cooking recipes maintained by a Ph.D. student. Also, links made to websites may be created for different reasons than the citation of scholarly content: often links are part of directory listings, and are to an institution as a whole rather than to a specific research output. Nevertheless, it is worth noting that journal Impact Factors have similar methodological issues, but that does not stop their widespread (albeit controversial) use to compare journals, research groups and institutions.[10]

Potentially, the use of webometric analysis also enables macroscopic comparison to be made of the relationships between country domains. The data for links between

Australia, New Zealand and UK academic domains will be used to compare the extent to which two ex-dominions of Britain relate to each other in cyberspace. The hypothesis will be examined that the pattern of linking between universities is the same whether the source is the UK or other Australasian universities using WIFs and a new measure of the tendency of universities to interlink. It would be interesting if evidence were to be produced that this was not the case, for example if the universities with lower research quality were ignored, or, conversely, if research quality itself did not affect the number of links from distant universities. An examination of the propensity of universities to link to each other will also be undertaken on a country by country basis.

Methodology

Three sets of link data were obtained. A specialised web crawler was used to obtain link data from the main UK, Australian, and New Zealand Universities. In each country, the specialised web crawler was restricted to universities officially recognised as such, for instance in Australia those listed at the Australian Vice Chancellors Committee site (http://www.avcc.edu.au/). An alternative approach would be to crawl all sites in the academic domain (ac.uk; edu.au; ac.nz) however the logistics of achieving this for the specialised crawler would be prohibitive, and the inclusion of non-university sites would complicate the analysis. The result of these crawls was a set of lists of the number of links made from universities in the source countries to academic institutions in the target countries:

- UK to Australia;
- UK to New Zealand;
- Australia to New Zealand;
- New Zealand to Australia;
- Australia to Australia;
- New Zealand to New Zealand.

The crawl produced numbers of links to all academic institutions in the target countries, but for the current study only links to the recognised universities were used.

For the intra-country link counts (au to au and nz to nz), links were only counted *between* institutions, not links *within* institutions. In other words, external links were being counted. Internal links are problematic for two reasons: they are probably not a measure of the overall impact of the web site; and they are often expressed as relative links rather than absolute links, so that a commercial search engine query may not include them.[8]

In the current study the number of academic staff was determined from the World Higher Education Database 2000.[11] It should be noted, though, that there appear to be inconsistencies in how institutions report staff numbers, for instance in distinguishing part-time and general staff. In some cases data was also gathered from institutional web sites.

Baseline university research measures: the Research Quantum

In order to assess Web Impact Factor metrics, it would be useful to have a recognised measure of the research quality of institutions as a baseline for comparison. This has been achieved for the UK in other WIF studies through the use of data from the official government Research Assessment Exercise.[12,13] In the previous version of this paper[7] bibliometric measures were used as well as the AsiaWeek 2000[14] rankings. The latter were not available for all universities, however, only the more successful ones, but some significant correlations were found for the bibliometric measures. In this study we have chosen to use the official Australian government research infrastructure grants instead for Australia and to focus upon this country for the WIF section of the results. This funding, known as the Research Quantum (RQ), is broadly similar to the UK's Research Assessment Exercise in ultimate objective, but is calculated through a simple publicly available formula[15] rather than peer review. The formula is crude in some respects, for example including a simple count of journal articles irrespective of quality. It is dominated by a calculation based upon research funding received from other sources and so a more serious concern for our purposes is that universities tending to specialise in more expensive research would have a higher RQ value. These concerns aside, it is clear that in general terms higher RQ values reflect more research at an institution. The value that will be used as an approximation to the average quality of a university's research is its RQ divided by the number of academic staff members. In fact this is perhaps more of a measure of the average quantity of research than its average quality, but we believe that the two are related. Bond private university has been excluded from this study as a result of it being ineligible for an RQ grant.

The specialised crawler and coverage issues

The specialist information science web crawler used has been described in detail previously.[16] It was designed to be able to conduct very precise crawls, in terms of accurately identifying duplicate pages with different URLs, and in not using heuristics to decide when to stop crawling, but to attempt to cover sites in entirety. This link

following program was started at the home page of each site covered unless it did not contain any references to other pages, in which case a seed page containing a list of all the university's departments was identified and used instead. In all cases except Liverpool University, for reasons described below, this seemed to produce substantial coverage of the target site.

It is important to include a discussion of the coverage of web sites obtainable by crawlers: both specialist ones and those used to populate commercial search engine indexes. For any crawler to visit and record a web page, the following key criteria must be met.

1. It must be able to find the URL of the page. The specialist web crawler could have found the URL from a hyperlink in a page previously indexed, but a commercial search engine could also follow links found on other web sites, and retrieve URLs found in previous crawls or submitted by users. Pages not linked to would not be found by the specialist crawler and would be less likely to be found by search engine crawlers.[9] It is not possible to be able to calculate the percentage of pages actually found by a crawl unless some alternative reliable method of page counting is available.

2. It must have permission to access the page. Some pages are only accessible to computers inside the institution (Intranet), or are password or IP protected (Extranet), or are specifically banned from being retrieved by crawlers (the "robots.txt" convention). Liverpool University's site was banned from being indexed by the specialist crawler.

3. Its internal algorithm must allow it to retrieve the page. The crawler may not be able to retrieve pages consisting of program-generated HTML, as opposed to static web pages. Program-generated pages are sometimes flagged by data at the end of the URL, indicated by a question mark (e.g. search.cgi?keyword=information).

4. Its internal algorithm must not reject the page, once retrieved. Rejection may happen if it is adjudged to be sufficiently similar to a previously downloaded page, or its HTML cannot be interpreted. Partial rejection is also possible for large documents, although this was not a criteria used by the crawler used for this research.

5. The host web server must be operating. Some web servers may be switched on only occasionally, others may have been decommissioned or out of service for a long period of time. Commercial crawlers should be able to avoid some problems by rescheduling repeat download attempts over long periods of time.

Probably the main differences between a specialist crawler and that of a big search engine like AltaVista are that the latter can crawl more of an academic web site by using its list of previously known URLs to start the crawl, but is less successful at identifying duplicate pages.[12]

It is important to emphasise that any crawl of a large academic web site is likely to be incomplete, as alluded to in 1-5 above. The results of this paper, therefore, refer to the subset of the web sites that were actually covered in each case. The underlying assumption is that this partial coverage is still a meaningful object of study and capable of producing significant results. This has been previously demonstrated for the UK.[12,13]

Search syntax for AltaVista

As a comparison with the crawler results, links between the universities were counted using the advanced text mode of the AltaVista search engine (http://www.AltaVista.com/sites/search/textadv).

In using commercial search engines such as Alta Vista, it has been found in the past that the results returned are inconsistent (for instance equivalent Boolean search statements may return differing results). This appears to be due to the priorities placed by the commercial search engines on returning useful results quickly, rather than precise results at any cost of processing time. At the time of the current study the inconsistencies between times of searching, and between different formulations of Boolean search statements, were negligible. While in past studies[7,17] has been felt necessary to carry out searches at different times and using a range of Boolean formulations, in the current study the consistency of the search engine was sufficiently good that single searches were carried out.

The search statements used for the AltaVista results were of the form (where xxx is the third level domain name of the target institution, e.g. Australian National University=anu):

UK-Australia	host:ac.uk AND link:xxx.edu.au
UK-New Zealand	host:ac.uk AND link:xxx.ac.nz
Australia-New Zealand	host:edu.au AND link:xxx.ac.nz
New Zealand-Australia	host:ac.nz AND link:xxx.edu.au
Australia-Australia	host:edu.au AND NOT host:xxx.edu.au AND link:xxx.edu.au
New Zealand-New Zealand	host:ac.nz AND NOT host:xxx.ac.nz AND link:xxx.ac.nz

While these gave results that were broadly comparable with the crawler results, it should be noted that the AltaVista results use all academic sites for the source of links (ac.uk; edu.au; ac.nz), the crawler only uses recognised universities as its source.

Boolean searches could be created on AltaVista that would only use the recognised universities, for example for NZ-Australia links:

(host:auckland.ac.nz OR host:aut.ac.nz OR host:ait.ac.nz
OR host:waikato.ac.nz OR host:massey.ac.nz
OR host:vuw.ac.nz OR host:canterbury.ac.nz OR
host:lincoln.ac.nz OR host:otago.ac.nz) AND
link:xxx.edu.au

These would become very complex and probably lead to a greater degree of inconsistency in the search engine results. As an experiment searches of this format were carried out for the NZ-NZ links, and were found to be very similar to those that used all of the ac.nz domain as the source.

Search syntax for AllTheWeb

The searches using AltaVista were carried out in May 2001. In July 2001 we became aware that AllTheWeb (otherwise known as FAST, http://www.alltheweb.com/) had introduced search syntax comparable with AltaVista for carrying out webometric analysis. In addition, AllTheWeb has been reported as having more consistent retrieval of link data.[18] In August 2001 the searches which had been carried out using AltaVista were repeated using AllTheWeb.

The search statements used for the AllTheWeb results were of the form:

UK-Australia	url.host:ac.uk +link.all:xxx.edu.au
UK-New Zealand	url.host:ac.uk +link.all:xxx.ac.nz
Australia-New Zealand	url.host:edu.au +link.all:xxx.ac.nz
New Zealand-Australia	url.host:ac.nz +link.all:xxx.edu.au
Australia-Australia	url.host:edu.au -url.host:xxx.edu.au +link.all:xxx.edu.au
New Zealand-New Zealand	url.host:ac.nz -url.host:xxx.ac.nz +link.all:xxx.ac.nz

Results

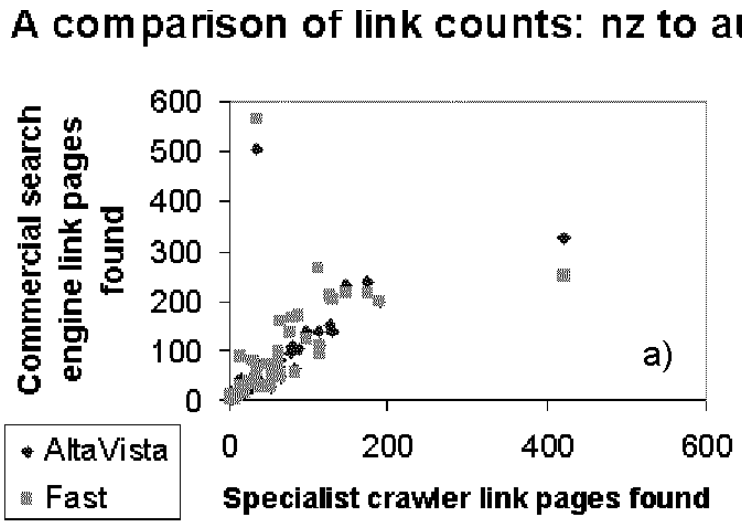
A comparison of link counts between methods

The main objective of this paper is to compare different national sources of links rather than different counting methods, but a brief prior consideration of this issue is presented here to address the issue of the reliability of the figures obtained. The three sources of link counts were tested at different times over a one-year period and so, given

the expanding nature of the web, the results could be expected to be considerably different from each other. The differing crawling and counting algorithms used would also add an extra source of variation. The issue of reliability is, therefore, of relative rather than absolute magnitude of the counts presented and so the three sources will be compared pair-wise, with a strongly linear trend tending to support the hypothesis that any conclusions drawn from the data about national differences are not artefacts of the collection mechanism.

If counts of links *from* a single academic web site were compared over time then the changes observed could be expected to include large jumps. Big increases could occur, for example, if a new large collection of pages were to be added and big decreases if sets of old pages were to be periodically deleted. Counts of links to the same site should be steadier, at least in relative size, depending as they do upon the contents of pages from a range of other sites. An exception to this occurs when one university introduces or removes a collection of pages with a disproportionately large collection of external links. This happens when a university hosts a copy of a collection of pages created at another institution and the originators have included a credit link on each page. Anomalies are therefore to be expected in any comparison of sources of link counts over time even if there is an otherwise linear trend.

Figures 1a,b,c and 2a,b,c compare sources of links pair-wise, for a different originating and target institution each, using the specialist crawler as the reference set.



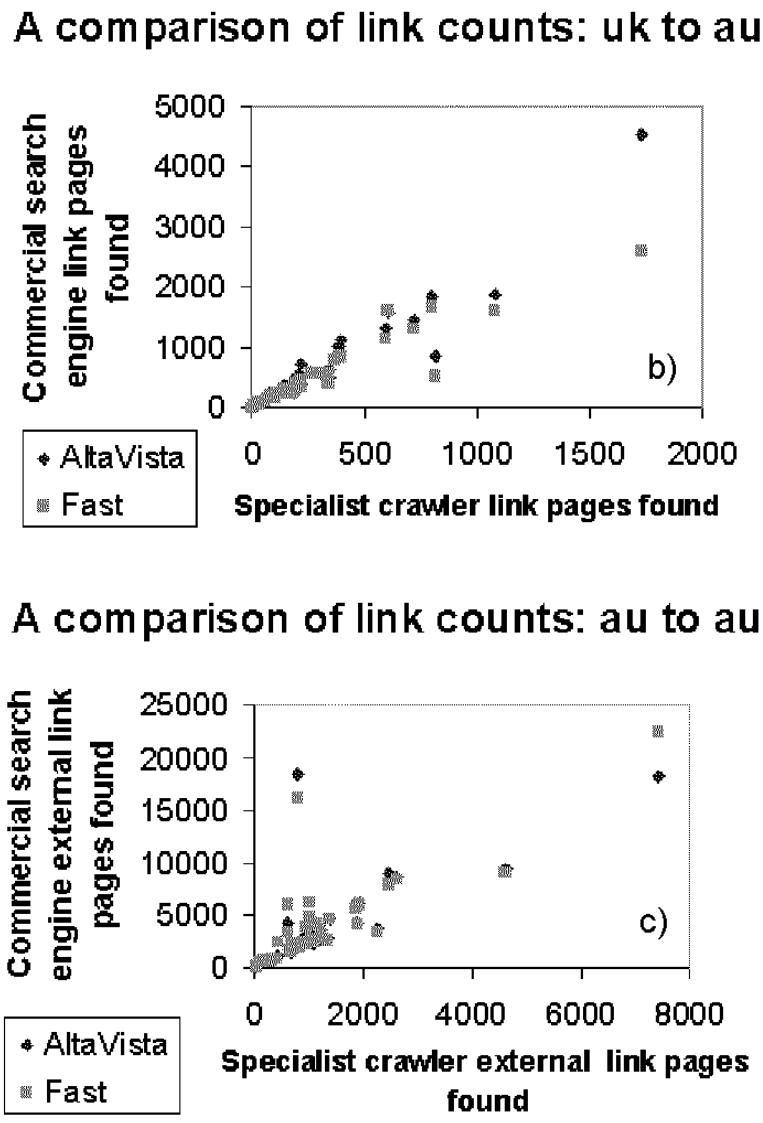
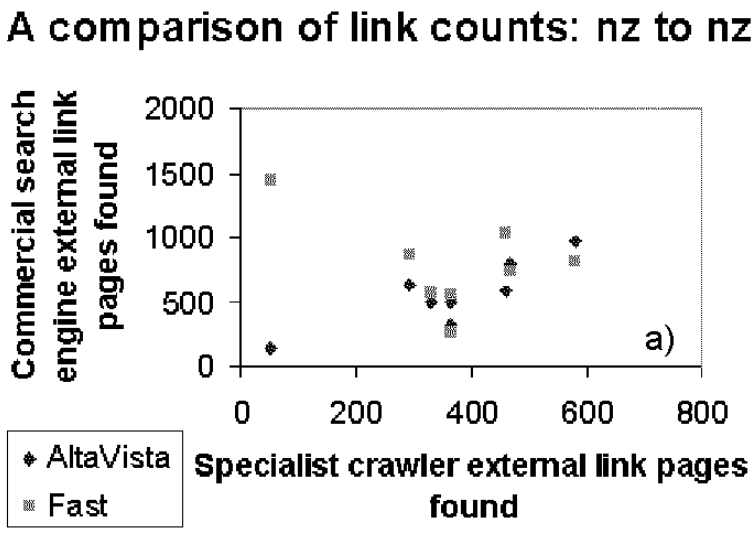
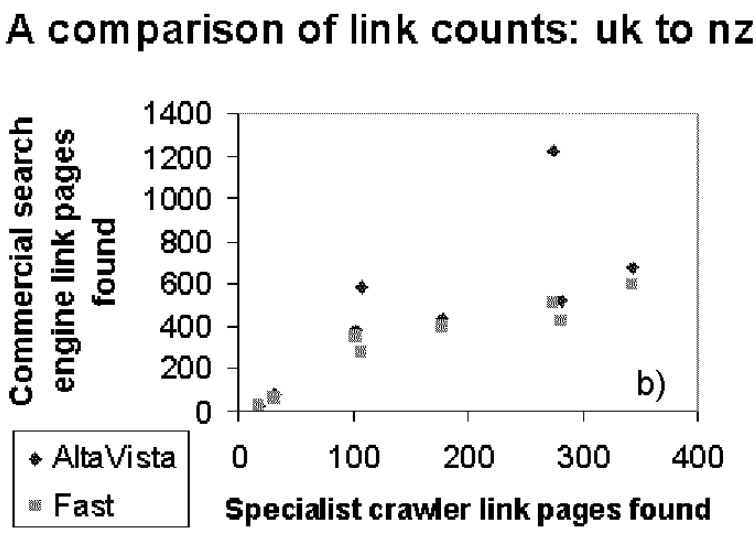


Figure 1. Comparisons of search engine counts to Australian universities with specialist crawler counts. Specialist crawler: July, 2000, AltaVista: May 2001, AllTheWeb: August 2001.


A comparison of link counts: nz to nz
Commercial search engine external link pages found
2000
1500
1000
500
0
0
200
400
600
800
a)
AltaVista
Fast
Specialist crawler external link pages found


A comparison of link counts: uk to nz
Commercial search engine link pages found
1400
1200
1000
800
600
400
200
0
0
100
200
300
400
b)
AltaVista
Fast
Specialist crawler link pages found

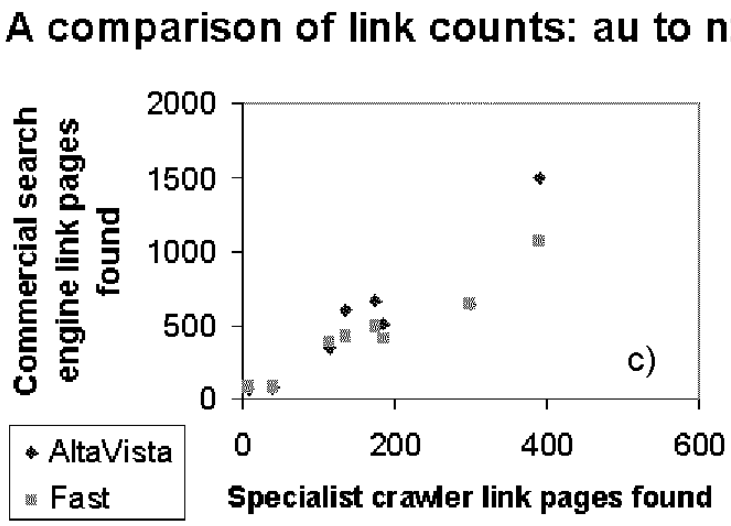


Figure 2. Comparisons of search engine counts to New Zealand universities with specialist crawler counts. Specialist crawler: July, 2000, AltaVista: May 2001, AllTheWeb: August 2001.

The graphs must be examined carefully to fully interpret their meaning because in many places two points overlap and only one can actually be seen. From observations of the graphs, several points can be noted.

- Linear trends are clear from all of the graphs except perhaps Figure 2a, with the commercial crawlers showing close agreement in most cases, both being proportionally two to three times greater than the specialist crawler.
- AltaVista tends to report the same or more links than AllTheWeb, except for the case of NZ-NZ.
- There is clear scatter around trend line, plus outliers in Figures 1a and 1c.

In conclusion, there is some evidence that data from different sources may be comparable in relative magnitude, but caution must be exercised because of the exceptions.

Average tendency to link to Australian and New Zealand universities

Table 1 shows the total links from universities by country, as found by the specialist crawler. Only external links are included, so that links from a university to itself are

omitted. These figures are not particularly helpful because there are different numbers and sizes of universities in the countries. In order to normalise the data to make the figures directly comparable, a measure of overall size of the source and target universities must be used as a denominator. For this, the total number of academic staff in all the universities concerned was used. This was used rather than web site page counts, because site page counts are a particularly unreliable figure.[12] The product of total staff for both source and target countries was used as the denominator. The overall result was modified by multiplying by n/(n-1), where n is the total number of universities, in the case where the source and target universities were the same. This was to compensate for universities links to themselves being excluded. The results of this calculation were divided by the highest resulting score (NZ to NZ: 0.000,113) to give a figure that we call the *normalised propensity to link*.

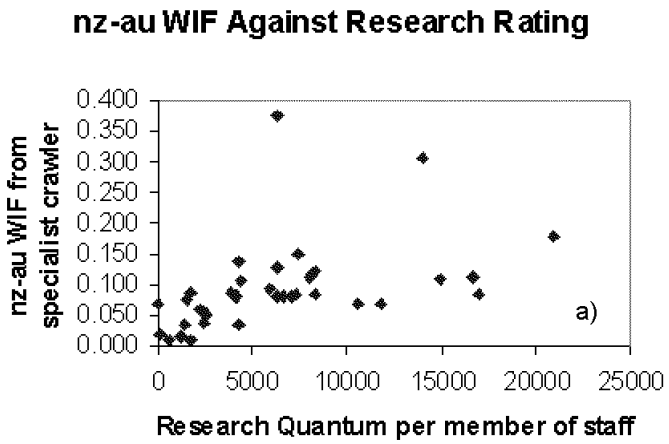
Link direction	Total links	Normalised propensity to link
NZ to NZ	2,909	1.00
AU to AU	47,128	0.52
NZ to AU	2,742	0.16
AU to NZ	1,342	0.08
UK to AU	11,512	0.04
UK to NZ	1,333	0.02

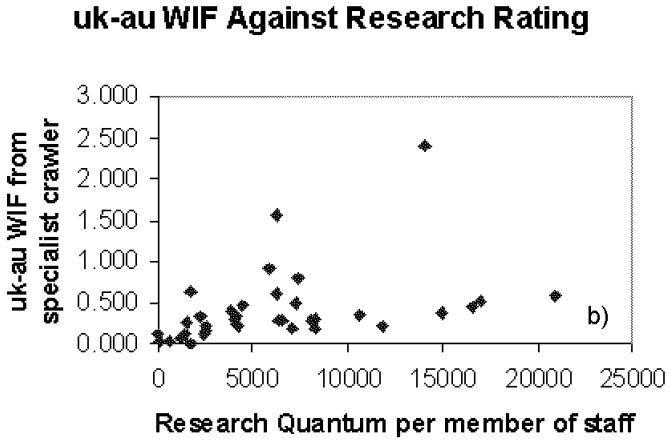
This shows total external links between universities, as found by the specialist crawler in July, 2000. It also shows this figure divided by the total number of staff in the crawled source institutions, and also by the total number of staff in the target institutions. For AU to AU and NZ to NZ the result of this is also multiplied by n/(n-1) where n is the number of universities crawled in the country.

From this table it can be seen that countries link most to themselves, which is to be expected. It also suggests that Australasian universities link to each other better than the UK does, with Australia being a more popular target of links in both cases. This suggests that the geographic proximity of Australia and New Zealand has resulted in better web connections. The results overall suggest that the academic web in New Zealand is somewhat insular; relatively well interconnected, but less well known internationally. Alternative explanations are possible, however, with countries other than the UK perhaps linking better to New Zealand. In chemistry, for example, co-authorship measures showed the UK connected better to Australia than New Zealand, with the latter connecting best to Australia and Sweden.[19] Interestingly, however, in *Glänzel*'s[20]

map of international scientific co-authorship for 1996/97, New Zealand is one of the most isolated of the advanced nations, connected only to Australia, with Australia also connected to the UK and the USA. Connections in *Glänzel*'s map indicated a degree of international co-authorship above a minimum threshold.

A comparison of counts of web links with the baseline research rating





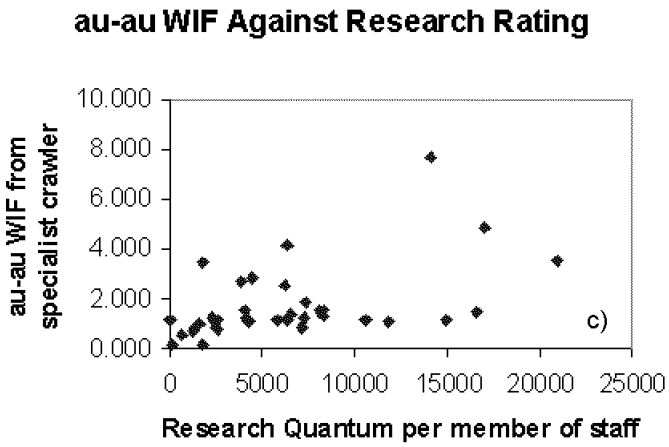


Figure 3. WIFs for Australian universities calculated from links from a) New Zealand, B) British, c) other Australian universities, plotted against Research Quantum per member of staff

Figures 3a-c show that (academic staff-size denominator) WIFs calculated from other universities in New Zealand, the United Kingdom and Australia all display some correlation (Pearson: 0.459, 0.401, 0.498 respectively with n=37, significant at the 1%, 5% and 1% levels, respectively) with the RQ-derived research measure. The graphs show wide scattering, however, showing that the relationship is far from simple. The WIFs calculated in the original paper were[7] were for the combined set of all Australasian universities ranked by Asiaweek. These were compared with citations per member of staff, publications per member of staff and Asiaweek rank, with only citations providing a significant correlation, at the 5% level, although Asiaweek rank did correlate at the 5% level with the original page count denominator WIF.

The scattering exhibited in the graphs appears to be too great to be attributed to problems with the RQ. It suggests that WIFs are influenced by other factors, such as the design of the target web site, the use of computing in the target university and the existence of a few highly linked to pages[12] that make the distribution of links to a site highly skew and therefore subject to larger variations than might be expected by chance.

Comparing the graphs with each other, the degree of scattering makes it too difficult to draw general conclusions about the distance factor in link creation. There is certainly no evidence to suggest that the less prestigious universities are being ignored by their peers or overseas when it comes to being the target of hyperlinks. In fact the graphs are consistent with the hypothesis that the quality of research at a university is the primary

determining factor for links. This is despite the fact that the majority of links may not be directly or indirectly research related, as was the case in one analysis of types of UK-to-UK links.[12] In order to fully understand these results, further results to establish a plausible explanation for this is needed.

Discussion

Both specialised crawlers and general-purpose search engines are useful in webometric studies. This study indicates that they broadly produce similar results. In general AltaVista and AllTheWeb found more links than the specialised crawler. However this may not be because of better coverage, but could also be due to the time factor, outdated or duplicate links being included in the search or the inclusion of non-university academic web sites.

In summary, specialised crawlers;

- Have known characteristics;
- Can be directed at specific domains;
- Require setting up and configuration.

While search engines:

- Use proprietary algorithms which are not known to the researcher;
- Cover a wide range of domains, but it is not clear to what depth they index a given domain;
- Are easily accessible and usable.

It is likely that future webometric studies will use both types of tools. Search engines are useful for exploratory studies, and as a "reality check" on a specialised crawler, while specialised crawlers should be able to produce more precise, targeted data.

While the links between the UK, Australia and NZ are fairly predictable, there might be potential for studying the links made by universities farther afield, for example to explore if Australian universities have more links to Asia because of their relative proximity. Indeed a web version of *Glänzel*'s[20] international co-authorship map would be a very interesting source of comparison.

The current study has revealed an issue in assessing university domains, that university domain names are not consistent or constant. For instance during the period of the study Auckland Institute of Technology changed to the Auckland University of Technology, with a consequent change of domain name from ait.ac.nz to aut.ac.nz.

However, site maintainers do not change links immediately, and the methodological issue is the extent to which obsolete domain names are included in searches for research oriented links. In the current study links to both names were taken into account (e.g., in AltaVista, link:ait.ac.nz OR link:aut.ac.nz). One drawback of using AllTheWeb was that the use of multiple brackets in search statements did not work consistently. For instance, in order to measure links from other NZ institutions to the Auckland University of Technology it was necessary to use the statement:

url.host:ac.nz -(url.host:aut.ac.nz url.host:ait.ac.nz)
+(link.all:ait.ac.nz link.all:aut.ac.nz)

This produced 4088 pages. However separate searches for ac.nz pages linking to ait.ac.nz and aut.ac.nz produced a total of only 1447 pages, even though this included links between ait.ac.nz and aut.ac.nz that should have been excluded by the combined statement. For the purpose of the current study, the separate searches were used. In contrast, AltaVista did produce consistent results for the equivalent statement.

Another issue is the devolution of research functions to separate entities that may have their own domain names. For instance both the Christchurch Medical School (chmeds.ac.nz) and the Wellington Medical School (wnmeds.ac.nz) are organisationally part of the University of Otago (otago.ac.nz) and arguably should be included for the purpose of link analysis, yet discovering the identities of *all* web sites affiliated with any given university is a non-trivial task. The Australasian Legal Information Institute AustLII (austlii.edu.au) is a joint venture of the University of Technology Sydney and the University of New South Wales, but was not counted as part of either institutions link count. However the specialised web crawler found 6363 pages linking to austlii.edu.au, almost twice as many as the total number of links made to UTS and UNSW.

Conclusions

One objective of this paper was to ascertain whether there was any evidence for differing patterns of link creation to Australasian universities from UK universities and between themselves. The propensity to link between universities showed national links were much more likely than international links. New Zealand also appeared to be relatively insular, interlinking well, but being less well linked to than Australia, an ironic conclusion given the authorship of this paper. UK institutions created about a quarter of the density of links to Australian and New Zealand universities then they did to each

other, evidence of a regional affinity. The data was not clear enough to reveal any more specific pattern that could be linked to the type of universities considered. If such a factor was present, then it was obscured by the wide scattering of the data as a result of the skewness of the distribution of links to pages.

Despite the fact that the WIFs calculated were significant in the sense of correlating with research quality they proved not to be a sufficiently reliable tool to give a convincing answer to the question of identifying real differences between the sources of links. Further work is perhaps needed to deal with the issue of data skewness in WIFs in order to produce a more powerful tool.

References

1. R. R. LARSON, Bibliometrics of the world wide web: An exploratory analysis of the intellectual structure of cyberspace, In *Proceedings of the 1996 Annual ASIS Meeting*, Baltimore (1996) pp. 71–79.
2. T. C. ALMIND, P. INGWERSEN, Informetric analyses on the World Wide Web: methodological approaches to "Webometrics", *Journal of Documentation*, 53 (4) (1997) 404–426.
3. R. ROUSSEAU, Sitations: an exploratory study, *Cybermetrics*, 1 (1997). Available: http://www.cindoc.csic.es/cybermetrics/articles/v1i1p1.html
4. P. INGWERSEN, Web Impact Factors, *Journal of Documentation*, 54 (2) (1998) 236–243.
5. E. GARFIELD, The impact factor, *Current Contents*, (1994) June 20. Available: http://www.isinet.com/isi/hot/essays/journalcitationreports/7.html
6. M. THELWALL, Web Impact Factors and search engine coverage, *Journal of Documentation*, 56 (2) (2000) 185–189.
7. A. G. SMITH, M. THELWALL, Web Impact Factors and university research links, In: *Proceedings of the 8th International Conference on Scientometrics and Informetrics*, Vol. 2 (2001) pp. 657–664, Sydney: Bibliometric and Informetric Research Group, UNSW.
8. A. G. SMITH, A tale of two web spaces: comparing sites using web impact factors, *Journal of Documentation*, 55 (5) (1999) 577–592.
9. S. LAWRENCE, C. L. GILES, Accessibility of information on the Web. *Nature*, 400 (1999) 107–109.
10. J. GOWRISHANKAR, P. DIVAKAR, M. BAYLIS, M. GRAVENOR, R. KAO, Sprucing up one's Impact Factor (two letters to the editor), *Nature*, 401 (1999) 321–322.
11. WORLD HIGHER EDUCATION DATABASE 2000 (2000) London: Macmillan Reference. [CD-ROM]
12. M. THELWALL, Extracting macroscopic information from the web, *Journal of the American Society for Information Science and Technology*, (2002, to appear).
13. M. THELWALL, Sources of links for academic Web Impact Factor calculations, *Journal of Documentation*, (2002, to appear).
14. ASIAWEEK, Asia's best universities 2000, (2000). Available: http://www.asiaweek.com/asiaweek/features/universities2000/rank_country/
15. DETYA, Higher Education Report for the 2000 to 2002 Triennium, (2001). Available: http://www.detya.gov.au/highered/he_report/2000_2002/html/3_8.htm
16. M. THELWALL, A Web crawler design for data mining, *Journal of Information Science*, 27 (2001) 319–325.

17. R. ROUSSEAU, Daily time series of common single word searches in AltaVista and NorthernLight, *Cybermetrics*, 2/3 (1998/9). Available: http://www.cindoc.csic.es/cybermetrics/articles/v2i1p2.html
18. J. BAR-ILAN, How much information the search engines disclose on the links to a web page? - A case study of the Cybermetrics home page, In: *Proceedings of the 8th International Conference on Scientometrics & Informetrics, Sydney 16-20 July 2001* (2001) pp. 63–73. Sydney: Bibliometric and Informetric Research Group, UNSW.
19. W. GLÄNZEL, A. SCHUBERT, Double effort = double impact? A critical view at international co-authorship in chemistry, *Scientometrics*, 50 (2) (2001) 199–214.
20. W. GLÄNZEL, National characteristics in international scientific co-authorship relations, *Scientometrics*, 51 (1) (2001) 69–115.

Received October 12, 2001.

Address for correspondence:
ALASTAIR SMITH
School of Communications and Information Management
Victoria University of Wellington
Aotearoa/New Zealand
E-mail: Alastair.Smith@vuw.ac.nz