The Web Impact Factor: A Critical Review [*]

Alireza Noruzi
Department of Library and Information Science, University of Tehran, Tehran, Iran
Email: anouruzi@yahoo.com

Abstract

Purpose – We analyse the link-based web site impact measure known as the Web Impact Factor (WIF). It is a quantitative tool for evaluating and ranking web sites, top-level domains and sub-domains. We also discuss the WIF's advantages and disadvantages, data collection problems, and validity and reliability of WIF results.

Design/methodology/approach – A key to webometric studies has been the use of large-scale search engines, such as Yahoo and AltaVista that allow measurements to be made of the total number of pages in a web site and the total number of backlinks to the web site. These search engines provide similar possibilities for the investigation of links between web sites/pages to those provided by the academic journals citation databases from the Institute of Scientific Information (ISI). But the content of the Web is not of the same nature and quality as the databases maintained by the ISI.

Findings – This paper reviews how the WIF has been developed and applied. It has been suggested that Web Impact Factors can be calculated as a way of comparing the attractiveness of web sites or domains on the Web. It is concluded that, while the WIF is arguably useful for quantitative intra-country comparison, application beyond this (i.e., to inter-country assessment) has little value.

Originality/value – The paper attempts to make a critical review over literature on the WIF and associated indicators.

Keywords: Webometrics, Link analysis, Bibliometrics, Journal Impact Factor
Paper type: Literature review

Background of the Study

Bibliometric research methodologies of library and information science have always been used to provide tools for understanding the dynamics of disciplines, developing policy, and justifying research funding. Since 1996 increasing efforts have been made to investigate the Web, as a significant scholarly medium for science and scholarship, by applying bibliometric techniques (Larson, 1996; Turnbull, 1996). Term applied to this new area of study include "webometrics" (Almind & Ingwersen, 1997). Webometrics is defined as: "the study of the quantitative aspects of the construction and use of information resources, structures and technologies on the Web, drawing on bibliometric and informetric approaches" (Björneborn, 2004).

The idea of measuring average link frequencies, that is, the WIF, as one of the quantitative indicators was developed in 1998 by Peter Ingwersen. Note that prior to Ingwersen, Rodríguez i Gairín (1997) had introduced the concept of information impact on the Internet in a Spanish documentation journal, but his article was not as influential as Ingwersen's. The WIF was based on an analogy between hyperlinks and citations and was the adaptation of the journal "*Impact Factor*" (originally proposed by Garfield in 1972) for the Web. However, the time periods for the WIF and the journal *Impact Factor* (JIF) are different. The JIF measures citations made in journals published during one time period to articles published in another

---

[*] Noruzi, A. (2006). **The Web Impact Factor: a critical review**. *The Electronic Library*, 24.

time period while the WIF is a "*snapshot*" of a search engine database at a specific time. Compared with the content of a journal paper, the content of a web resource lacks peer review and thus lacks quality control. The WIF is therefore not exactly the equivalent of the JIF. However, the WIF was inspired by the JIF.

Egghe (2000) pointed out that hyperlinks are very different from citations. They can be synchronic (web pages can link to each other regardless of their publication date) while citations are diachronic. Generally, only previously published papers can be cited by later published ones, not vice versa. However, it is also possible for authors to cite each other's paper at the same time due to publication delays and the existence of invisible colleges.

There has been much recent interest in analysing links and web pages, especially with commercial search engines, like AltaVista (Rodríguez i Gairín, 1997; Ingwersen, 1998; Smith, 1999; Snyder & Rosenbaum, 1999; Thelwall, 2000; Thelwall, 2001; Smith & Thelwall, 2002; Thelwall, 2002; Kousha & Horri, 2004; Noruzi, 2005), AlltheWeb and Yahoo, despite the relatively unregulated and problematical nature of this information source. In fact, the number of link pages returned by Yahoo or AltaVista are actually the number of web pages that contain at least one link to the web site in question.

Information professionals have studied the idea of utilizing bibliometric and informetric methodologies to the Web, and have started to lay the basis for a newly emerging area of Webometrics. The novel aspect is to regard the Web as a citation network where traditional information entities and citations from them are replaced by web pages (e-articles) and hyperlinks respectively. In this context, these resources are the entities of information on the Web, with hyperlinks from them acting as citations.

Webometric techniques are still in their experimental stage in testing whether the classical bibliometric methods applied to the Web are reliable and feasible means of comparing and analysing web sites.

Definition

The WIF provides quantitative tools for ranking, evaluating, categorizing, and comparing web sites, top-level domains and sub-domains. There are three types of link. Outgoing links from web pages are here named outlinks, links coming into a site from other sites inlinks (backlinks), and links within the same site (from one page to another page) self-links. And there are three types of WIF: overall WIF, inlink (revised) WIF, and self-link WIF. For the overall WIF of a web site, the numerator is the number of inlink pages from outside the site and self-link pages within the site (see Figure 1); for the inlink (revised) WIF, the numerator is the number of inlink pages counted from outside the site (see Figure 2); and for the self-link WIF, the numerator is the number of self-link pages counted from within the site. The denominators all remain the same, the number of web pages within the web site in question.

Figure 1: Calculation for Web Impact Factor

$$\mathbf{WIF} = \frac{\mathbf{A} = \text{total link pages (all inlink and self-link pages)}}{\mathbf{D} = \text{number of web pages published in the web site which are indexed by the search engine, not all web pages available in the web site}}$$

WIF= A/D = Web Impact Factor

The WIF is a form of measurement used to determine the relative standing of web sites in particular fields, or a country; for instance, academic web sites in a country. The higher the impact factor, the higher the perceived reputation of the web site. The WIF answers the question "*what impact has this web site had?*". A WIF is a measure of the frequency with which the "average web page" in a web site has been linked at a given point in time. In

general, a web site with a higher impact factor may be considered to be more prestigious or of a higher quality than those web sites with a lower impact factor.

The WIF gives a measure of average (external or absolute) impact per page, which could be for a single university web site or all web sites in an entire country, for example. This has a close analogy with impact factors for journals, as, indeed, link analysis has with citation analysis generally. As stated above, the WIF is naturally different from the JIF. Given the increasing use of WIFs –as well as the (less explicit) use of web site prestige– in evaluation, a critical examination of this indicator seems necessary. Informed and careful use of web impact data is essential. Users may be tempted to jump to ill-formed conclusions based on WIF statistics unless several caveats are considered.

The WIF is useful in clarifying the significance of inlink (or total link) frequencies. It eliminates some of the bias of such counts which favour large web sites over small ones, or well-known web sites over less-known ones, and of older web sites over newer ones. Particularly in the latter case such web sites have a larger linkable body of literature than smaller or younger web sites. All things being equal, the larger the number of previously published web pages, the more often a web site will be linked. Thus the greater the number of link pages to a web site, the greater the WIF will be. The symbolic role played by the links in representing the content of a web site is an extensive dimension of information retrieval, and can expand the scope of information seeking by retrieving not only those web sites that have linked a primordial web site, but also those that related to the linking web sites.

Revised Web Impact Factors
Researchers can establish analogous impact factors, (excluding self-links), for the web sites they are evaluating. Self-links often represent an important percentage of the links that the pages of a web site receive.

Self-links reflect the logical structures used for organising web pages in the local servers (Ingwersen, 1998). The WIF analysis for self-links is less meaningful than inlinks, because the majority of self-links within a web site can be created for navigation purposes rather than for endorsing the content of target pages (Smith, 1999; Thelwall, 2000). The bigger the web site the larger the number of self-link pages will tend to be.

Inlinks from outside represent more effort to point to target pages, and thus contain more valuable information. However, it is not always easy to separate self-links from inlinks (Li, 2003). For the revised WIF calculation, see Figure 2.

Figure 2: Calculation for WIF revised to exclude self-links

A= total links to a web site (all inlink and self-link pages)

B= inlinks to the web site (this is a subset of A)

C= self-links and navigational links within the same web site

D= number of web pages published in the web site which are indexed by the search engine, not all web pages available in the web site

R-WIF= revised WIF (**B** / **D**)

Link Search Strategies
The Yahoo search engine can be used to count the number of links and web pages for webometric research because it has advanced search facilities. This search engine continuously records links as represented by web pages from a large number of the world's web sites. Yahoo has "*link*" and "*linkdomain*" commands for link searching, but in the current

research, we use only the "*linkdomain*" command. For retrieving total links, inlinks, or self-links to the web site of the *University of Tehran*, as an example, we can use the following syntax in Yahoo (see Figure 3).

Figure 3: Lack of standardization for link search strategies in Yahoo

Total links:	No. of hits
1- *linkdomain:*ut.ac.ir/ *OR linkdomain:*www.ut.ac.ir/	17000
2- *linkdomain:*www.ut.ac.ir/ *OR linkdomain:*ut.ac.ir/	16700

Inlinks	No. of hits
1-a. *linkdomain:*ut.ac.ir/ *NOT domain:*www.ut.ac.ir/	13100
1-b. *linkdomain:*ut.ac.ir/ *NOT host:*www.ut.ac.ir/	16500
1-c. *linkdomain:*ut.ac.ir/ *-host:*www.ut.ac.ir/	4940
2-a. *linkdomain:*ut.ac.ir/ *NOT host:*ut.ac.ir/	16600
2-b. *linkdomain:*ut.ac.ir/ *-host:*ut.ac.ir/	4940
3-a. *linkdomain:*www.ut.ac.ir/ *NOT host:*www.ut.ac.ir/	9030
3-b. *linkdomain:*www.ut.ac.ir/ *-host:*www.ut.ac.ir/	2680
4-a. *linkdomain:*www.ut.ac.ir/ *NOT host:*ut.ac.ir/	8670
4-b. *linkdomain:*www.ut.ac.ir/ *-host:*ut.ac.ir/	2680
5-a. *linkdomain:*ut.ac.ir/ *OR linkdomain:*www.ut.ac.ir/) *NOT* (*host:*ut.ac.ir/ *OR host:*www.ut.ac.ir/)	4930
5-b. (*linkdomain:*ut.ac.ir/ *OR linkdomain:*www.ut.ac.ir/) *NOT host:*www.ut.ac.ir/	16800
5-c. (*linkdomain:*ut.ac.ir/ *OR linkdomain:*www.ut.ac.ir/) *NOT host:*ut.ac.ir/	16500
6-a. (*linkdomain:*ut.ac.ir/ *OR linkdomain:*www.ut.ac.ir/) *NOT* (*domain:*ut.ac.ir/ *OR domain:*www.ut.ac.ir/)	7800
6-b. (*linkdomain:*ut.ac.ir/ *OR linkdomain:*www.ut.ac.ir/) *NOT domain:*ut.ac.ir/	7800
6-c. *linkdomain:*ut.ac.ir/ *NOT domain:*ut.ac.ir/	7740

Self-links	No. of hits
1-a. (*linkdomain:*ut.ac.ir/ *OR linkdomain:*www.ut.ac.ir/) *AND* (*host:*ut.ac.ir/ *OR host:*www.ut.ac.ir/)	163
1-b. (*linkdomain:*www.ut.ac.ir/ *OR linkdomain:*ut.ac.ir/) *AND* (*host:*ut.ac.ir/ *OR host:*www.ut.ac.ir/)	164
1-c. (*linkdomain:*ut.ac.ir/ *OR linkdomain:*www.ut.ac.ir/) *AND* (*host:*www.ut.ac.ir/ *OR host:*ut.ac.ir/)	164
2-a. (*linkdomain:*ut.ac.ir/ *OR linkdomain:*www.ut.ac.ir/) *AND host:*ut.ac.ir/	164
2-b. (*linkdomain:*ut.ac.ir/ *OR linkdomain:*www.ut.ac.ir/) *+host:*ut.ac.ir/	164

3-a. (*linkdomain:*www.ut.ac.ir/ *OR linkdomain:*ut.ac.ir/) *AND host:*ut.ac.ir/	164
3-b. (*linkdomain:*www.ut.ac.ir/ *OR linkdomain:*ut.ac.ir/) +*host:*ut.ac.ir/	163
4-a. (*linkdomain:*ut.ac.ir/ *OR link:*www.ut.ac.ir/) *AND host:*www.ut.ac.ir/	163
4-b. (*linkdomain:*ut.ac.ir/ *OR link:*www.ut.ac.ir/) +*host:*www.ut.ac.ir/	164
4-c. (*linkdomain:*ut.ac.ir/ *OR link:*www.ut.ac.ir/) +*domain:*www.ut.ac.ir/	302

As shown in Figure 3, the syntax of the "*linkdomain:*" command in combining with other commands such as "*host:*" or "*domain:*" is not standard in search engines at this time (September 7, 2005). Figure 3 shows that the link search strategies of Yahoo are unorganised and anarchic. In addition, we should note that the outcome from Yahoo applying the inverted but logically identical set operations actually differs slightly. Consequently, it seems that the retrieved results are not reliable, and the number of web pages indexed and reported for a particular web site by Yahoo is not complete. There are considerable differences between reported links and retrievable links using Yahoo.

Li et al. (2003) used AltaVista for data collection and confirmed that to calculate the WIF with the number of web pages in the site as denominator is not reliable. Noruzi (2005) has done a time series WIF for the University of Tehran in order to monitor AltaVista search engine performance, and shows that there exist large increases and decreases during the time.

It is evident that there is no standardization for link search strategies in Yahoo, and therefore, there is no standardization for retrieving linkers. This is also true for other search engines, like AltaVista, as we tested the listed commands in Figure 3 on AltaVista during a time series. In some cases, assessment of target pages showed that they have been retrieved because of email addresses with searched domain name. In other words, using the "*linkdomain:*" or "*link:*" commands in Yahoo may retrieve resources with email addresses related to the searched domain.

WIF Advantages and Utilities
Each method has its advantages and disadvantages, including the WIF. The major advantages and utilities of the WIF include the following:

1. The WIF analysis method presents a methodology for evaluating "*international visibility*" and impact of institutional and academic web sites, as well as their competitive relations to other web sites. The WIF can be regarded as a useful tool to measure the relative visibility of a company, organization, or country on the Web. It must be noted that the WIF is not the only indicator of the use, visibility, and popularity of a web site.

2. The WIF provides a way to evaluate a web site's relative importance, especially when we compare it to others in the same field or a country's domains. Therefore, to compare web sites we should stick to a particular category. We do not compare web sites in different research fields. So, the WIF measures the success and relative influence of similar web sites.

3. The WIFs of national, sector, and larger web segments or top-level domains are calculable. Comparisons should be performed with caution, and preferably be carried out within the same snapshot, and also comparisons should be limited to comparable units. The variation of the WIF over different snapshots taken within short intervals does exist (Ingwersen, 1998; Wormell, 2001).

4. The WIF provides a quantitative indicator of web sites long-term influence. In the final analysis, impact simply reflects the ability of web sites and webmasters to attract users and

cybercitizens, and consequently backlinks. We nevertheless warn against the indiscriminate use of WIF data.

   5. The WIF may in turn provide novel insights into the retrieval process on the Web. For instance, clusters of web sites can be detected by means of link page co-occurrence. Moreover, The WIF can be regarded as a tool for measuring the *accuracy* of web search engine performance and web site organisation, linking, and structuring of pages (Ingwersen, 1998; Wormell, 2001).

WIF Misuses and Incorrect Applications

The WIFs can be influenced and biased (intentionally or otherwise) by many factors and variables. However, when using them for evaluation it is important to realize that WIFs are influenced by other factors, such as the following:

   1. The WIF depends on the search engines' web coverage. The coverage varies considerably between countries. Therefore, the impact factor of any web site will be proportional to the search engines' coverage. Furthermore, the web site sets in search engines' databases are not constant but may vary in composition from month to month. So, the results of analyses of the Web by search engines can only be regarded as rough indications rather than definite conclusions. As stated by Ingwersen (1998) "in principle almost all searches will thus be incomplete." It should be noted that search engines data are inherently incomplete.

   2. Assume that there are only two web sites in an area of interest: web site A and web site B. Assume that web site A was linked 10 times and published 10 web resources over the time period being examined. Similarly, web site B was linked 100 times and published 100 web resources over the same time period. Both web sites have the same impact factor, but can it be said that the two web sites have had the same influence on the literature? It would be appropriate to suggest that both web sites are equally efficient in attracting links, but web site B has had a greater contribution to the current literature by a factor of 10 times. Therefore, large web sites that publish many web resources may not have as high an impact as smaller scientific and research web sites, because the high web resources rate counteracts the high inlinks rate (Noruzi, 2005).

   3. Web page inlink rates determine the web site impact factor, but not vice versa. In other words, the impact factor of a web site is not statistically representative of its web pages. The WIF is a function of the number of inlinks per web site not per web page. For the web site's impact factor to be reasonably representative of its web pages, the inlink rate of individual web pages in the web site should show a narrow distribution. Assigning the same score (the WIF) to all web pages masks this tremendous difference –which is the exact opposite of what an evaluation is meant to achieve. Inlinks frequencies for individual web pages are quite varied. Even the unlinked web pages are then given full credit for the impact of the few highly linked web pages that predominantly determine the value of the WIF.

   4. The international language bias extends further than simply the language of publication. English-speaking webmasters and authors rarely link non-English language literature, regardless of the otherwise implicit appropriateness of the link itself. Several earlier publications have raised this concern and discussed its indirectly negative effect on the impact factor standing of non-English language web sites. This problem has been compounded by the fact that webmasters from non-English-speaking environments seek to publish their works in English. For example, Japanese web sites may not be as highly linked to due to language differences. "One may note that the current Japanese WIF is far below the expected mean value for countries and sectors. This situation leads to considerations about the influence of language as well as of national cultural and social factors on the meaning and interpretation of WIFs in general" (Ingwersen, 1998). One might argue that English-speaking webmasters may not link to the Japanese, Chinese, or Persian web sites, but today the use of English

documents means that few important resources are missed. According to Li (2003) English is the dominant language on the Web in Western Europe. Countries sharing the same language tend to link more than those that do not. Web sites in the English language may have higher impact factors.

Search engines' databases have an English language bias. They are dominated by English web sites. The preference of the search engines' databases for English language web sites will contribute to a low WIF for the few non-English web sites that are included, since most links to web sites in languages other than English are given by other web sites in the same language. Links in the national language of the web site are preferred by the web site's webmasters and web bloggers. Previous studies have shown that sites tend to link more within their own country than outside (Bharat *et al.*, 2001; Thelwall, 2002; Vaughan & Thelwall, 2004).

5. The original intention for the use of the WIF was to allow comparison between the link rates of web sites. The application of this tool evolved into a means with which to assess the quality of the web sites themselves, on the basis of the premise that a higher rate of link indicated higher web site quality. Furthermore, the misuse of this calculation has, in recent years, widened to include evaluation of the quality of world university web sites ranking, it is not meaningful, it is dangerous. The frequency of link has been adopted as a rough indicator of quality. However, a high link rate may not always be associated with high quality. In addressing the extension of this tool to academic evaluation, while the WIF may provide a gross approximation of the prestige of web sites, it does not advise using this value as the sole means of comparative evaluation. It should be apparent that the ranking relationship between quality and link is not absolute. We can mainly judge the quantity and quality of web sites. Quantity is easily evaluated, involving counting the number of web pages, whereas quality is a notoriously difficult aspect to appraise, in that subjectivity and bias frequently overshadow the process. Evaluating web site quality is a notoriously difficult problem which has no standard solution. Even citations and journal Impact Factors can only measure the impact of research rather than its quality (Seglen, 1997; Moed, 2002).

6. Relations between fields strongly determine the WIF. Therefore, interdisciplinary web sites attract more inlinks and web sites in the fields with a high turnover of research or developing technology usually have larger impact factors. Extension of the WIF to cross-discipline web site comparison is also inappropriate. Different disciplines and specialties exhibit a different impact factor range. In this way, web sites may be viewed in the context of their specific field. It is rare to find that the ranking of a web site will change significantly unless the web site's influence has indeed changed. The WIF is dependent on the research field, topics that are currently more trendy are therefore more likely to be linked. Many factors contribute to the number of inlinks per web site, including the science; the nature of the research; the style of the communication; language; readership and diffusion (more readers equals more inlinks); conformism (webmasters often link those web sites that are currently linked). Smith (1999) found for Australasian universities no significant correlation between research output, measured in terms of publication counts, and the equivalent of the AltaVista original general WIF. Smith and Thelwall (2002) and Thelwall (2001) showed that WIFs correlated significantly with the research rating and quantity. Li (2003) stated that the significant correlation coefficients found between WIFs and research ratings support their use.

7. Unlike scientific citations to journals, institutions or individuals, which may be stable or may constantly increase, the number of pages linking up to a particular web resource may indeed decrease or disappear over time, for example, due to closedowns or restructuring of web sites. Thus, in contrast to the common citation Impact Factor calculation a retrospective WIF is not reproducible (Ingwersen, 1998). Therefore, the WIF depends on dynamics (expansion or contraction) of the web site.

8. In comparison between the WIF and the JIF, it is clear that some journals have been accused of trying to manipulate their impact factors by asking the authors to increase the number of references to papers published in them (Smith, 1997). It is a distortion of the scientific process, and also some webmasters are asking others to link to their sites. It should also be pointed out that should backlink analyses be used for studies that have a real impact upon universities, they would be extremely vulnerable to manipulation. This stems from the unrefereed nature of the medium (Web) and from the ease by which large numbers of web pages can be generated automatically by those with programming skills (Thelwall, 2001).

9. Web sites introducing a new technique, a new computer programme, a new model, a new test, diagnostic or outcome criteria, reviews and controversies, also garner their share of links, and these vary among areas and with different stages of development of a discipline. The WIF rankings within a field or a country are more meaningful than between fields or countries. The WIFs have some value as relative measures in closely related defined categories but do not have validity as absolute or relative measures across categories.

Discussion

A web site's ultimate success depends upon its quality, distribution, and many other competitive factors including its size, language, age, visibility, and popularity and not only its impact factor. The WIF should not be used without careful attention to the many phenomena that influence inlink rates, as for example the link creation motivation. The WIF should be used with other web indicators. There are many artifacts that can influence a web site's impact and its ranking in search engines, not the least of which is the inclusion of the authority pages, because they attract more inlinks. In the WIF analysis, a good authority is a well-linked, popular page on a topic.

A wide range of factors can influence the WIF. Ease of access to web sites, publication immediacy, site language, site updating and type of publication material (for example publishing an e-journal) have all been identified as contributors to the WIF. Scientific articles tend to link only scientific articles. In a similar context, general web sites tend to have higher impact factors than specialist web sites, because of the larger pool for link. Controversial topics may increase impact. Nothing will replace the judgment necessary for webmasters to publish high quality resources.

Overall, the WIF should always be used in combination with other indicators when evaluating countries or institutions. For nation by nation comparisons, there is very little controversy about the use of link indicators. Comparisons of web sites on the basis of their impact factors should be limited solely to intra-area evaluation; we warned that inter-area comparisons may be both inappropriate and misleading. In addition, some web sites in certain fields or topics require more time to mature because of delayed recognition or because of the time required to produce research documents or experimental results. Invidious comparisons between web sites even in different fields do not take these subtleties into account. Successful webmasters know that in order to improve the quality of web sites, there is no substitute for judgment, quality, and relevance. Impact and other link measures merely report the facts.

It is necessary that the WIF data must be checked in a relatively short period of time to avoid creating bias due to links disappearing as the web pages were taken down or moved to a different location (Thelwall, 2001). Those who choose to use the WIF as a comparative tool should be aware of the nature and premise of its derivation and also of its inherent flaws and practical limitations.

However, direct comparison between web sites on the basis of the total number of links alone is an inappropriate measure influenced by a number of factors. Thus, a relative adjusted score was sought to allow direct quantitative comparison. Interpretation of the WIF must be done with an understanding of its true meaning, that it reflects the number of times, on

average, that a source item in a particular web site is likely to be linked. Clearly, while the use of WIFs to compare web sites may be arguably possible on a qualitative level, it cannot be done on a quantitative basis.

Conclusions

Evaluation of web sites is a formidable but necessary task considering the wide range of choices available. The WIF, as explained in the above, is a useful tool for evaluation of web sites, but it must be used discreetly. Considerations include the amount of web pages or other types of material published in a web site, contents, and variations between disciplines. The web site's status in regard to coverage in the search engines' databases as well as the occurrence of a domain name change is also very important. The WIFs are always approximate and not absolute. The WIF of a site is not stable, because everyday some webmasters are deleting the old outlinks to several web sites and others are linking to new ones. The WIF would still be far from being a quality indicator: link impact is primarily a measure of scientific utility rather than of scientific quality. For evaluation of scientific quality, there seems to be no alternative to qualified experts reading the web site resources. All WIF studies should be normalized to take into account variables such as field, or discipline, country, language, and link practices.

The WIF is not a perfect tool to measure the quality of web sites but there is nothing better and it has the advantage of already being in existence and is, therefore, a technique for quantitative evaluation of web sites. Despite the recognition that the WIF is an imperfect measure and seven years of criticism, there is no obvious alternative. Thus, those forced to use this tool for direct web site comparison should be encouraged to remain open-minded and cautious, with an awareness of the inherent limitations of its use. While the WIF is arguably useful for quantitative intra-country comparison, application beyond this (i.e., to inter-country assessment) has little value. In the future, there may be more sophisticated ways of assessing the quality of web sites.

Acknowledgements

References

Almind, T.C., & Ingwersen, P. (1997), "Informetric analyses on the World Wide Web: methodological approaches to Webometrics", *Journal of Documentation,* Vol. 53 No. 4, pp. 404-426.

Bharat, K., Chang, B.W., Henzinger, M., & Ruhl, M. (2001), "Who links to whom: mining linkage between web sites", In: *Proceedings of the IEEE international conference on data mining (ICDM)*, San Jose, November 2001, pp. 51-58.

Björneborn, L. (2004), "Small-world link structures across an academic web space: a library and information science approach", PhD dissertation. Copenhagen: Department of Information Studies, Royal School of Library and Information Science.

Egghe, L. (2000), "New informetric aspects of the Internet: some reflections-many problems", *Journal of Information Science*, Vol. 26 No. 5, pp. 329-35.

Garfield, E. (1972), "Citation analysis as a tool in journal evaluation", *Science*, Vol. 178 No. 4060, pp. 471-9.

Ingwersen, P. (1998), "The calculation of web impact factors", *Journal of Documentation*, Vol. 54 No. 2, pp. 236-243.

Kousha, K., & Horri, A. (2004), "The relationship between scholarly publishing and the counts of academic inlinks to Iranian university web sites: exploring academic link

creation motivations", *Journal of Information Management and Scientometrics*, Vol. 1 No. 2, pp. 13-22.

Larson, R.R. (1996), "Bibliometrics of the World Wide Web: an exploratory analysis of the intellectual structure of cyberspace", In: Hardin, S. (Ed.), *Proceedings of the 59th Annual Meeting, ASIS 96*, Baltimore, pp. 71–79.

Li, X. (2003), "A review of the development and application of the web impact factor", *Online Information Review*, Vol. 27 No. 6, pp. 407-417.

Li, X., Thelwall, M., Musgrove, P., & Wilkinson, D. (2003), "The relationship between the WIFs or inlinks of computer science departments in UK and their RAE ratings or research productivities in 2001", *Scientometrics*, Vol. 57 No. 2, pp. 239-255.

Moed, H.F. (2002), "The impact factors debate: the ISI's uses and limits", *Nature*, 415, 731-2.

Noruzi, A. (2005), "Web impact factors for Iranian universities", *Webology*, Vol. 2 No. 1, Article 11. http://www.webology.ir/2005/v2n1/a11.html

Rodríguez i Gairín, J.M (1997), "Valorando el impacto de la información en Internet: AltaVista, el "Citation Index" de la Red", [Impact assessment of information on the Internet: AltaVista, the Citation Index of the Web]. Revista Española De Documentación Científica, Vol. 20 No. 2, pp. 175-181.

Seglen, P.O. (1997), "Why the impact factor of journals should not be used for evaluating research", *British Medical Journal*, Vol. 314, pp. 498-502.

Smith, R. (1997), "Journal accused of manipulating impact factor", *British Medical Journal*, Vol. 314 No. 7079, p. 463.

Smith, A.G. (1999), "A tale of two web spaces: comparing sites using web impact factors", *Journal of Documentation*, Vol. 55 No. 5, pp. 577-592.

Smith, A.G., & Thelwall, M. (2002), "Web impact factors for Australasian universities", *Scientometrics*, Vol. 54 No. 3, pp. 363-380.

Snyder, H., & Rosenbaum, H. (1999), "Can search engines be used for web-link analysis? A critical review", *Journal of Documentation*, Vol. 55 No. 4, pp. 375-84.

Thelwall, M. (2000), "Web impact factors and search engine coverage", *Journal of Documentation*, Vol. 56 No. 2, pp. 185-189.

Thelwall, M. (2001), "Extracting macroscopic information from web links", *Journal of the American Society for Information Science and Technology*, Vol. 52 No. 13, pp. 1157-1168.

Thelwall, M. (2002), "Evidence for the existence of geographic trends in university web site interlinking", *Journal of Documentation*, Vol. 58 No. 5, pp. 563-574.

Turnbull, D. (1996), "Bibliometrics and the World Wide Web", (Technical Report FIS-12-19-1996-1). Faculty of Information Studies, University of Toronto. Available (October 15, 2005) http://www.ischool.utexas.edu/~donturn/research/bibweb.html

Vaughan, L., & Thelwall, M. (2004), "Search engine coverage bias: evidence and possible causes", *Information Processing & Management*, Vol. 40 No. 4, pp. 693-707.

Wormell, I. (2001), "Informetrics and webometrics for measuring impact, visibility and connectivity in science, politics and business", *Competitive Intelligence Review*, Vol. 12 No. 1, pp. 12- 23.