

The Relationship between the research activity of Iranian medical universities and their Web Impact Factor

Maryam Asadi*¹ - Maryam Shekofteh²

30 May 2008

Introduction

At present, the World Wide Web is one of the main information sources on scientific and search activities. It is, therefore, a suitable criterion for evaluating new communications methods in webmetrics. According to Kousha, the target of all webmetrics studies is to validate the web links as the new information resources and to evaluate their impacts on formal and informal relationships. Consequently, some researchers consider the similarities between citations and web-links. For instance, Rousseau uses the term “Situation” to refer to a used site. Ingwersen proposed “Web impact factor” by analogy with Journal impact factor (JIF) . Borgman &

Furner discussed on the similarities between “linking and citing”; Vaughan & Shaw suggested web citations counts can potentially be “the appendix or alternative ISI citation counts and as a module for impact factor.

University web sites in Iran, as in many countries, are large multifaceted communication devices, and are increasingly used for a wide variety of purposes, from introducing the universities, collages, departments and faculty members to student admissions, academic courses, library catalogues, offering research findings, accessing web sites, etc. The calculation of the web impact factor (WIF) is suggested as a tool for evaluating and identifying the level of university website impacts; this calculation is a means for

¹-Master of library and Information Science. Sharif University of Technolgy. Central library.e-mail: masadi1355@yahoo.com

² - Shahid Beheshti University(of Medical Science & Human Services), Department of Medical library & Information Science; PhD student of Library and Information science, Islamic Azad University(Science & Research Branch); . e-mail: shekofteh_m@yahoo.com

ranking of the universities' effective presence in the web. It should be noted that higher WIF reveals higher visibility and more success of the university websites at national and international levels. The idea of WIF was first initiated by Gairin (1977) in an article entitled "The Value and effect of information on the internet" in the "Spanish Journal of Documentations". A year later, Ingwersen (1998) discussed the matter in an English language journal, and therefore, provided the grounds for expanding the idea even further. Bearing in mind the importance of university web sites, and considering the fact that such web sites are evolved more than the other web sites, numerous researches have been conducted to show the links patterns among the universities (Rajabalibeglou, 2006; Zeynolabedini, et al., 2006; and Osareh, 2006 among others). Moreover, the relationships between research activities and link patterns have been considered as well (Chen et al, 1998; Thelwall, 2001a; Thelwall, 2002a; Thelwall, 2001b; Thelwall, 2002b; Tang and Thelwall, 2002). The first researchers were not actually looking for a significant correlation between web links and university research rankings; these researches were trying to analyze the unsafe nature and technical problems of citation studies in the web as well as the insufficiencies of the web search engines (Thomas & Willet, 2000; Chu & Thelwall, 2002).

There have been new methods to conduct citation analyses in the web. There have been some investigations on citation studies in e-journals, while some others have focused on the evaluation of the WIF, Domain citation analysis at national or organizational level, and link creation motivations (Kousha, 2003; Noruzi, 2005;

Kousha, 2006; Osareh, et al. 2007; Wilkinson et al, 2003; Noruzi, 2005).

The first part of the present research seeks to determine the rankings of the Iranian medical university web sites with regards to the total number of current pages on web sites, the number of self-links, and in-links. The study, then, evaluates the universities' web sites rankings with regards to WIF based on the number of the web pages as well as based on the number of the academic staff as its denominator. In another section of the study, the correlations between the two calculated WIF and research productivities as measured. For this purpose, each university ranking based on its published articles in journals which indexed by the Institute for Scientific Information (ISI) in 2006 are determined, then, the amount of its relations to the two calculated WIF are studied.

Method

To study the web impact factor, an appropriate search engine selection is a must. According to Smith, 1999, the search engine should not only cover an enormous part of the web and more hyperlinks, but it has to have advanced search facilities to count the links possessed by the web site as well. In previous studies, the researchers found out that Alta Vista search engine possesses the above criteria and is safer which very well covers the university web sites. Moreover, Alta Vista includes more hyperlinks compared to other search engines, and Alta Vista is usually used in webmetrics researches (Thelwall, 2001b; Li et al, 2003; Kousha & Horri, 2004; Noruzi, 2005; Vaughan & Thelwall, 2003). In the present study; therefore, Alta Vista search engine has been used to determine the number of pages of each university web site, and the number of in-links as well

as self-links. In our preliminary searches, utilizing Boolean Operators, different formulations were used which resulted in different outcomes. To reach certainty and to exactly determine a search strategy with the lowest degree of error, all search strategies from the available articles were extracted and preliminary searches were performed. The web sites searches were all performed in three days in November, 2007. Following consultations with the experts in the field, the following all search strategies were confirmed for searches in Alta Vista:

1) the number of pages at the web site:

domain:xxx.ac.ir OR domain:www.xxx.ac.ir;

2) the number of in-links:

(linkdomain:xxx.ac.ir OR linkdomain:www.xxx.ac.ir)

NOT (domain:xxx.ac.ir OR domain:www.xxx.ac.ir)

3) the number of self-links:

(linkdomain:xxx.ac.ir OR linkdomain:www.xxx.ac.ir)

AND (domain:xxx.ac.ir OR domain:www.xxx.ac.ir)

The following formula was used to calculate the WIF:

$$WIF = \frac{\text{Number of Web pages link to the target site (Inlink)}}{\text{Number of Web pages within the site}} \quad (1)$$

$$WIF = \frac{\text{Number of Web pages link to the target site (Inlink)}}{\text{Number of staff members associated with the site}} \quad (2)$$

In the second part of this research, the ISI Database was used to determine the number of published articles by the Iranian medical sciences universities in 2006 as it is one of the most important national criterion for measuring research productivities. This way, the rankings of the above medical sciences universities were determined. The number of the published articles was found through the advanced search part of Web of Science database with the following search strategy:

CU=iran AND OG=med* AND PY=2006

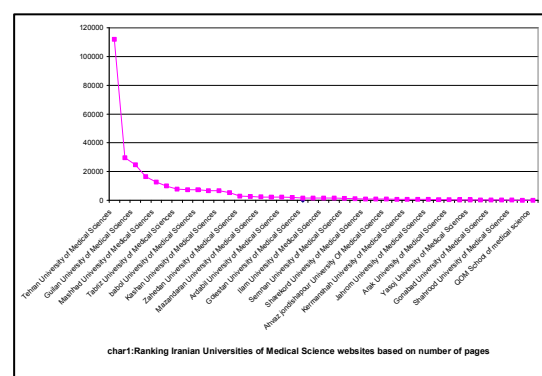
Finally, the amount of this relation to the rank with rank obtained from their WIF calculation was studied through the Spearman Correlation Coefficient.

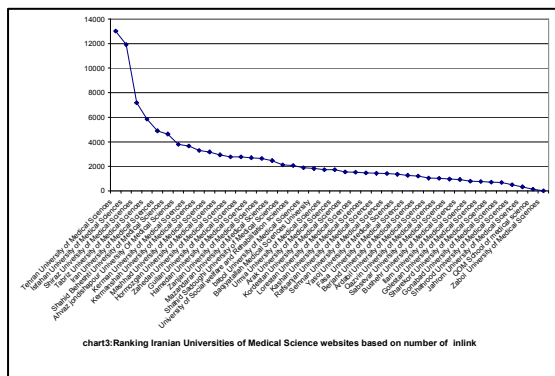
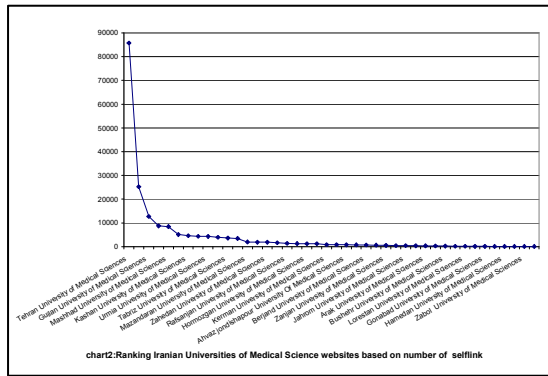
Data

The subjects of the present study included the medical universities. The list of the Iranian medical universities was obtained referring to the web site of the Iranian Ministry of Health and Medical Education. Then, the specific web addresses of each university were searched. It should be noted that some medical universities were not available in the list which were added to the final list (e.g. Baqiyattallah and Bojnood medical universities). The latter lacked any web site, so it was deleted from the research. Finally, 42 universities were selected as subjects of the study.

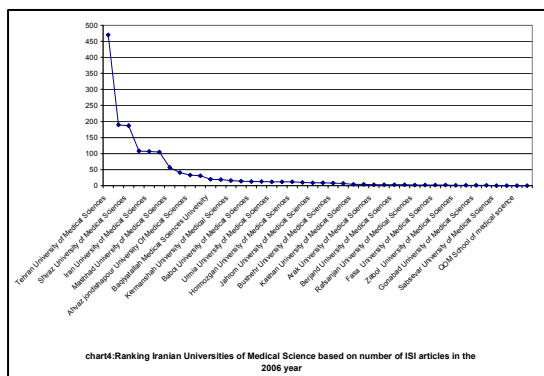
Results

The findings of the study on the ranking of the Iranian medical university websites concerning the number of the total pages in the website, the number of self-links and in-links revealed that Tehran, Iran, and Guilan medical universities obtained the highest ranks concerning the number of the available number of pages and self-links (Charts 1 & 2). Tehran, Isfahan, and Tabriz medical universities got the highest ranks of the number of the in-links (Chart 3).



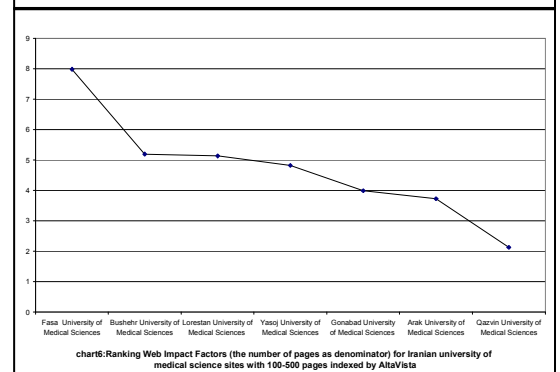
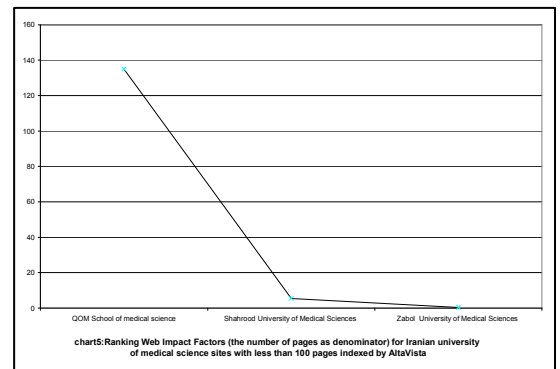


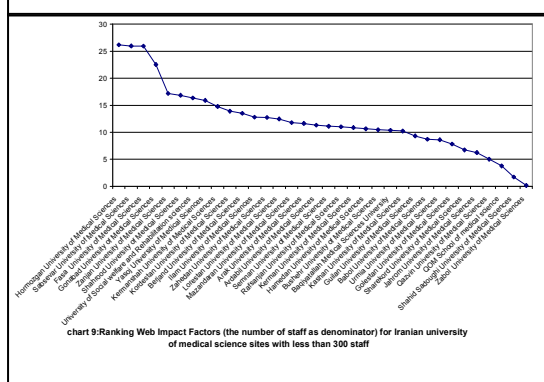
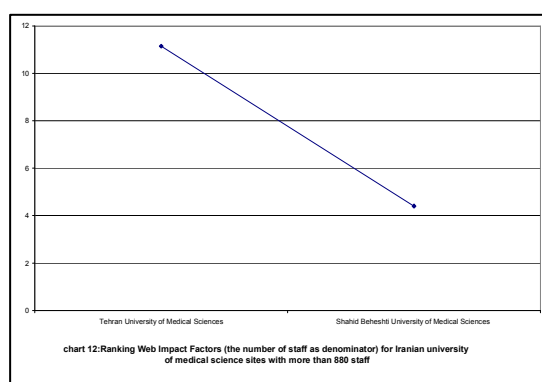
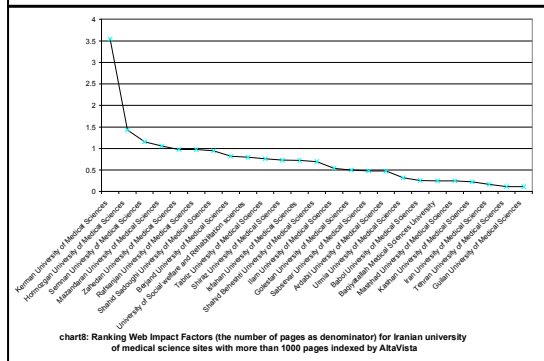
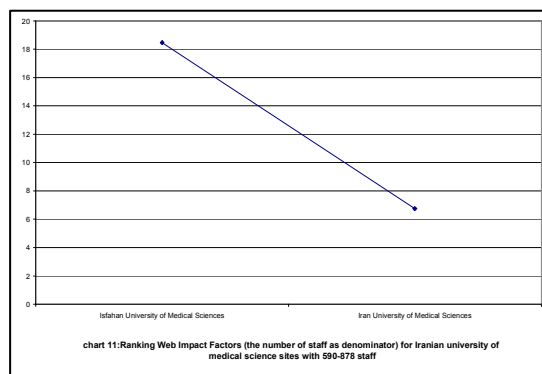
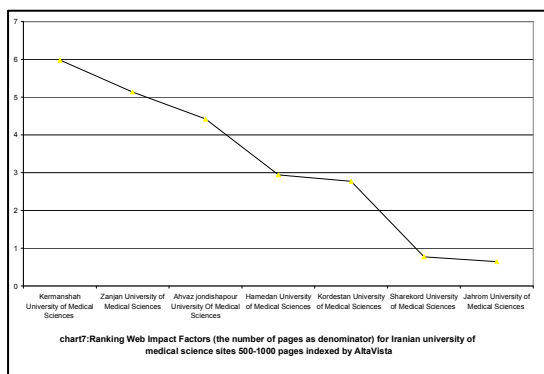
As it can be seen in Chart 4, Tehran, Shaheed Beheshti, and Shiraz medical universities have respectively obtained the highest ranks concerning the number of the published articles in ISI in 2006.



The rankings of the Iranian medical universities concerning WIF based on the number of the web pages were calculated in 4 groups. The findings showed that Kerman, Kermanshah, Fasa, and Qom medical universities obtained the highest WIF scores. Moreover, medical university WIFs concerning the number of the academic staff members (as the denominator) were calculated in 4 groups. The results revealed that Hormozgan, Shiraz, Isfahan, and Tehran respectively obtained the highest WIFs.

Charts 5-12 show the rankings of the medical universities' WIFs concerning the number of the available web pages on the web site and the number of the academic staff members (as the denominator), separately.





Question:

Is there a significant correlation between the ranking of the medical universities concerning the two calculated WIF (based on the number of the web pages and the number of the academic staff members as the denominator) and the research activities?

To study the correlations among the variables of the present research, the Spearman Correlation Coefficient was used as the frequency distribution showed a high skewness.

The study on the above correlation showed no significant relationship among the study variables; i.e., among the calculated WIF based on the number of pages and based on the number of the academic staff members (as the denominator), and the research activities (the number of the published articles in ISI), (Tables 1 & 2).

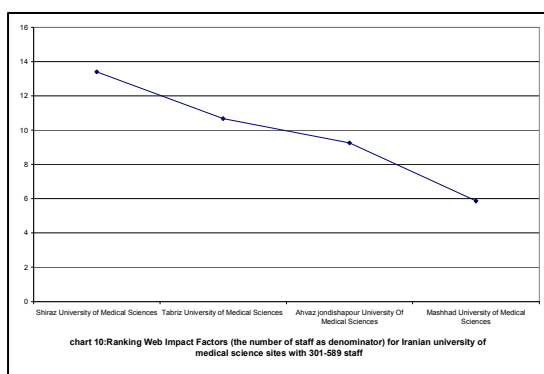


Table1- The relationships between WIF (the number of pages as denominator) and research activities

| | WIF (less than 100) | WIF (100-500) | WIF (500-1000) | WIF (more than 1000) |
|-----|---------------------|---------------|----------------|----------------------|
| ISI | -0.866 | 0.110 | 0.393 | 0.159 |
| N | 3 | 7 | 7 | 25 |

Table2- The relationships between WIF (the number of staff as denominator) and research activities

| | WIF (less than 300) | WIF (301-589) | WIF (590-878) | WIF (more than 880) |
|-----|---------------------|---------------|---------------|---------------------|
| ISI | -0.140 | 0.800 | 0.328 | 0.328 |
| N | 34 | 4 | 2 | 2 |

Discussion

The high self-links on a site is a mark of its correct link of information and resources. The users may find the information more easily through self-links. Moreover, search engines can provide better representations of the sites through self-links. In general, the amount of self-links is important in site retrieval and visibility (Osareh, 2007). The number of pages in a web site is a sign of its quantity. Moreover, the more the in-links, the more valid and important a site will be, and the better its visibility will be. The quantity of in-links in a web site determines its validity, importance, quality, visibility, and rank; it also shows that the information on the site is desired by the users.

In all the three above aspects, Tehran Medical University ranked the first. However, WIF analytical method is an approach to evaluate “international visibility” on the web. The WIF calculation showed that Qom medical University had the highest WIF score,

though it possesses only one web page (135), while Tehran Medical University with the highest number of web pages obtained a low WIF. The reason may lie in the formula to calculate WIF so that the websites with more web pages than their in-links, will show a lower WIF, and vice versa. This has been regarded as a drawback of the formula in other studies (Noruzi, 2005a; 2005b; and 2006). This drawback in the formula causes the websites with fewer numbers of pages to obtain lower WIF. This can cause problems in the process of calculating a website’s WIF and validity. The same problem can show itself in website’s academic staff members calculated WIF. Therefore, the results show that WIF index may not be an appropriate index for medical university website validation. It appears that the number of in-links is a better index for evaluation.

As it is mentioned in numerous studies, exact access to the number of available pages in the website through the search engines is bounded with some limitations including different retrieval algorithms. Therefore, web analysis results through the search engines can be only studied as *relative indexes rather than exact results*. As is stated by Ingwersen (1998), almost all searches are deficient (Noruzi, 2005b). It should be noted that the data and results obtained from the search engines is in nature deficient, so they have to be interpreted with caution. It is because our knowledge regarding the range and coverage of the search engines is low (Wilkinson et al., 2003). Bearing the above problems in mind, the authors faced problems in calculating the number of linked pages (in-links and self links), in the number of available web pages in the website, and in the links themselves to determine new orders. Following investigations on all orders from the previous studies (

Noruzi, Web Log; Kusha, 2005; Li, et al., 2003; Noruzi, 2005b) and consultations with experts, the advanced previous search order with the least error was used.

Web link analysis is an interesting outlook to evaluate the universities, though it is not a tool conduct a complete evaluation. Due to web dynamic nature, the search engine deficiencies (the data collection tools in such searches), lack of enough recognition of the motivations and causes of web link establishments, the presence of some other factors (except the links) in web evaluations, these studies and results in website rankings have to followed with caution.

Web experts have always sought to discover correlations among different continuous and discrete variables. The present study too, has tried to study and analyze the correlations between the WIF as a continuous variable and the other discrete variable (research activities).

Bearing the results obtained in mind, it seems that the web links are designed with the purpose of providing unofficial scientific and non-academic relations to guide and direct the users to access the Iranian medical universities' websites. Moreover, based on the previous studies (Kousha & Horri, 2004; Kousha, 2006) it seems that one of the important reasons for the low volume of the Iranian medical universities' web sites is, unlike other countries, the lack of web sites on the part of most academic departments, and consequently the academic staff members. Many Iranian universities have been unsuccessful at international levels to attract links for many reasons including the Persian language, lack of scientific information, lack of proper distribution of information in e-journals and e-conferences (Noruzi, 2005a).

There are many reasons to link to the websites.

Language, access to the news, important information resources, programs, information based on user needs, structural and content information in the site itself, the presence of full-text sources, presence of catalogs and union catalogs, etc., are all among sources to attract effective links. There are also reasons for the lack of inter-website links, such as language limitations, geographical, and political problems, the official relations among governments, social, cultural, ethnic, and racial problems, website technological problems, website address changes, changes in the website content and quality, website content weakness, lack of valid scientific information, and lack of proper distribution of the information in e-forms (Osareh, 2007). Yet, the weak presence of Iranian universities and research centres on the web environment, is not reasonable. the university officials and academic staff members to pay more careful attention to the websites under their direction, and convert the website to a valuable source for the students, academic staff members and clerks and for the whole Iranian society and the world; so bring them out of inactivity.

Suggestions

- Medical universities and colleges of medicine have to propose specific budgets for designing, and maintenance of their websites. They have to try to publish their websites in formats such as HTML, HTM, PDF, ASP, DOC, which the search engines provide a better coverage .
- There have to be a special space provided for all academic staff, clerks and students to be able to publish their scientific works, resumes, and career details. The academic staff members should be

required to offer their course details on the web site so that the students would be web-centred.

- The university web sites have to have a site map so the robots of the search engines can present them in the web.
- All journals published by the universities have to be presented on the web. This will increase their visibility and citations.

References

- Chen, C., et al. (1998). *How did university departments interweave the web: A study of connectivity and underlying factors. Interacting with Computers.* 10(4): 353-373.
- Chu, H.; Thelwall, (2002). *Library and information science schools in Canada and USA: A webmetric perspective. Journal of Education for Library and Information Science.*
- Kousha, K.(2005) *Webmetrics Approach for measuring citation Impact of open access LIS Journals. Informology .* 2(2): 185
- Kousha, Kayvan.(2004).*Comparing Iranian newspaper web sites using web impact factor. Informology. 1(2):88.*
- 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.1(2):13-22.*
<http://www.koosha.tripod.com/articles/scholarly.doc>
- Kousha, Kayvan (2006). *The relationship between scholarly publishing and the counts of academic inlinks to Iranian university web sites: exploring academic link creation motivations. Library and the Information Sciences: the Quarterly Journal of Organization*
- for Libraries, Museums and Documents Center of Astan-e Quds-e Razavi. 9 (2):35*
- Li, Xuemei and et al. (2003). *The relationship the WIFs or inlinks of Computer Science Departement in UK and their RAE ratings or research productivities in 2001. Scientometrics..57(2):*
- Noruzi A. (2005) *Web impact factor for Iranian Universities. Webology. 2(1) available at:< http://www.webology.ir/2005/v2n1/a11.html> [9/12/2007]*
- Noruzi, Alireza. *Knowledge Science weblog. Available in: www.nouruzi.blogfa.com.*
- Noruzi, Alireza (2004). *The Web Impact Factor: A Survey of Some Iranian University Web Sites. Studies in Education & Psychology, 5(2): 105-119.*
- Nouruzi, Alireza..(2006) *The web impact factor: a critical review. The Electronic Library:24 (4).*
- Osareh, Farideh (2006).*Collabration at national and international level amongst Library and Information Science School (LIS) websites..*
- Osareh, Farideh et al... (2007) . *Web links analysis in Library & information science international and national association and institute web sites. . Library and the Information Sciences: the Quarterly Journal of Organization for Libraries, Museums and Documents Center of Astan-e Quds-e Razav 10 (2): 52*
- Rajabalibeglou, Reza; Jokar, Abdolrasul (2006). *Study of the relationship between world university rankings and count of links to them: based on Shanghai university ranking and Times Higher Education. Informology. 4 (1,2): 179.*
- Smith, A. and Thelwall, M. (2002) *Web Impact Factors for Australasian universities, Scientometrics.. 54 (1-2).*

M. Kretschmer & F. Havemann (Eds.): Proceedings of WIS 2008, Berlin

Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting

Humboldt-Universität zu Berlin, Institute for Library and Information Science (IBI)

This is an Open Access document licensed under the Creative Commons License BY

<http://creativecommons.org/licenses/by/2.0/>

- Smith, A.G. *The Impact of Web sites: a comparison between Australasia and Latin America*. In: *Proceedings of INFO'99, Congreso Internacional de Informacion, Havana, 4-8 October (1999)*.
- Tang, R. and Thelwall, M. (2002) *Exploring the pattern of links between Chinese university web sites, Proceedings of the ASIST Annual Meeting, Volume 39*.
- Thelwall M., Harries G. (2003). *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*. *The American Society for Information Science and Technology*, 45(7):594-602. <<http://www3.interscience.wiley.com/cgi-bin/fulltext/104081845/HTMLSTAR1>> [9/13/2007]
- Thelwall, M. (2001a), *Results from a web impact factor crawler*. *Journal of Documentation*, 57(2): 177-191.
- Thelwall, M. (2001b) *A publicly accessible database of UK university website links and a discussion of the need for human intervention in web crawling*. University of Wolverhampton.
- Thelwall, M. (2002a), *A comparison of sources of links for academic Web Impact Factor calculations*. *Journal of Documentation*, 58(1), 66-78.
- Thelwall, M (2002b). *A comparison of sources of links for academic Web Impact Factor calculations*. *Journal of Documentation*, 58(1), 66-78.
- Thomas, O.; Willet, P. (2000) *Webmetric analysis of departments of librarianship and information science*. *Journal of Information Science*. 26 (6): 421- 428
- Vaughan, L., & Thelwall, M. (2003) *Scholarly use of the web: what are the key inducers of links to journal web sites? Journal of the American Society for Information Science and Technology* 54(1): 29-38.
- Wilkinson, David; Harries, Gareth; Thelwall; Mike & Price; Liz (2003) *Motivations for academic web site interlinking: evidence for the Web as a novel source of information on informal scholarly communication*. *Journal of Information Science* 29(1): 49-56.
- Zeynolabedini, Mohsen; Maktabifard, Lila and Osareh (2006). *Links analysis in World national library web sites*. *Studies in Education & Psychology*.