

Home	Table of Contents	Titles & Subject Index	Authors Index
----------------------	-----------------------------------	--	-------------------------------

Delinking: An Exploratory Study

[Isola Ajiferuke](#)

Faculty of Information and Media Studies, University of Western Ontario, London, Ontario, Canada N6A 5B7. E-mail: [iajiferu \(at\) uwo.ca](mailto:iajiferu@uwo.ca)

Received March 5, 2008; Accepted March 30, 2008

Abstract

The objective of this exploratory study is to determine the delinking practices of webmasters of colleges and universities in Canada and the US. An online questionnaire was created and all the 92 webmasters of Canadian colleges and universities were invited to complete it while the invitation was extended to only 500 of their counterparts in the US. A total of 123 webmasters responded to the survey with about 88 % of them indicating that they have had to remove external links from their web sites before. According to most of the webmasters, delinking is done as often as required but less than 20 times a year. Reasons for delinking a web site are categorized into obvious (e.g., broken/dead links, and content of linked site is no longer current or relevant) and unobvious (e.g., political/philosophical difference with the owner of the linked site, and request from the webmaster of the linked site). Measures to minimize delinking are discussed.

Keywords

Delinking; Link analysis; Universities; Colleges; Canada; US

Introduction

Informetrics is the quantitative study of information production, storage, retrieval, dissemination, and utilization ([Wolfram](#), 2000). The major areas of study within informetrics used to include: statistical aspects of language; characteristics of authors; characteristics of publication sources; citation and co-citation analysis; scientific productivity indicators; information growth and obsolescence; and document/information resource usage ([Tague-Sutcliffe](#), 1992). However, since the rise in popularity of the Internet, researchers have tried to extend the applications of informetrics methods to the World Wide Web. Such applications have often been referred to as webometrics ([Bjorneborn & Ingwersen](#), 2001).

A major thrust of webometrics studies is the analogy between a citation to a paper-based publication and an inlink (i.e. links pointing to or "sitation") to a web site ([Smith](#), 2004; [Vaughan & Shaw](#), 2003; [Prime, Bassecoulard & Zitt](#), 2002, [Rousseau](#), 1997). This has led some researchers to extend the comparisons to: motivations for citing and those for hyperlinking ([Kim](#), 2000; [Thelwall](#), 2003; [Wilkinson et al.](#), 2003; [Bar-Ilan](#), 2004; [Kousha & Horri](#), 2004); co-citations among scholarly publications and co-sitations among similar

web sites ([Larson](#), 1996; [Faba-Perez et al.](#), 2004; [Thelwall & Wilkinson](#), 2004); and journal impact factors and impact factors of web sites ([Ingwersen](#), 1998; [Smith](#), 1999; [Thelwall](#), 2000; [Vaughan & Hysen](#), 2002).

One major difference between citation and situation is that once a scholarly publication is cited, it cannot be uncited whereas a link to a web site can be removed. This phenomenon is known as delinking. The number of inlinks to a web site can increase the traffic to that site ([Lohse & Spiller](#), 2003) and has also been factored into the algorithms used by some search engines, such as Google to rank web pages retrieved in response to a query ([Brin & Page](#), 1998). Hence, delinking can decrease the traffic to a site as well as lower its rank in a set of pages retrieved by a search engine.

The main objective of this study is to investigate the extent of this phenomenon which has not attracted much attention from webometricians. Apart from a linked site being dead, we would like to find out what the other reasons for delinking are. For example, in 2002, the Authors Guild urged its members to remove links on their web sites to Amazon.com because it was offering used editions of current books ([Associated Press](#), 2002). It is not known if the members heeded this advice.

Literature Review

We shall be reviewing two sets of studies: those investigating the reasons for hyperlinking in the first place (before we then investigate why these links are later removed); and those investigating retraction of published articles, which is the closest phenomenon in print publication to delinking.

There have been many studies investigating the motivations for hyperlinking with most situated in the academic environment. These include: those classifying link creation motivations among schools in a discipline ([Chu](#), 2005), among university web sites in a country or region ([Thelwall](#), 2003; [Wilkinson et al.](#), 2003; [Bar-Ilan](#), 2004) or from universities in one country to those in another country ([Kousha & Horri](#), 2004); and those examining motivations for hyperlinking to open access journal articles or hybrid of research-related web sites ([Smith](#), 2004; [Kousha & Thelwall](#), 2006), or in electronic articles ([Kim](#), 2000). Content/context analysis was used to determine motivations for hyperlinking in all these studies except in the one by Kim where authors of electronic papers were interviewed to determine why they included links to external sites. Also, though these studies used different classification schemes for the hyperlink motivations, a common theme in all of them is that while a large percentage of the links were made for teaching or research reasons, a good percentage were also made for administrative/professional service and navigational purposes (e.g., link to resources, directories, subject indexes, personal home pages, etc.).

One hyperlink motivations study that was conducted outside the academic environment was by [Vaughan, Gao and Kipp](#) (2006). They examined links to 280 North American information technology companies, and listed the motivations for linking as acknowledging cooperate sponsors, link from customers, online directory, link from job advertisements, link to business partners, news articles about products or companies, list of products or clients or companies, and others (e.g. weblog, bookmark, product review, newsletter, etc.).

According to [Nath, Marcus and Druss](#) (2006), reasons for retracting articles could be categorized as misconduct (falsification, fabrication, or plagiarism) or unintentional error (mistakes in sampling, procedures, or data analysis; failure to reproduce findings; accidental omission of information about methods or data analysis). However, studies have shown that some retracted articles continued to be cited as valid work after being retracted

([Budd et al.](#), 1999; [Neale et al.](#), 2007), though at a reduced after retraction ([Garfield & Welljams-Dorof](#), 1990; [Pfeifer & Snodgrass](#), 1990; [Whitley et al.](#), 1994).

Methodology

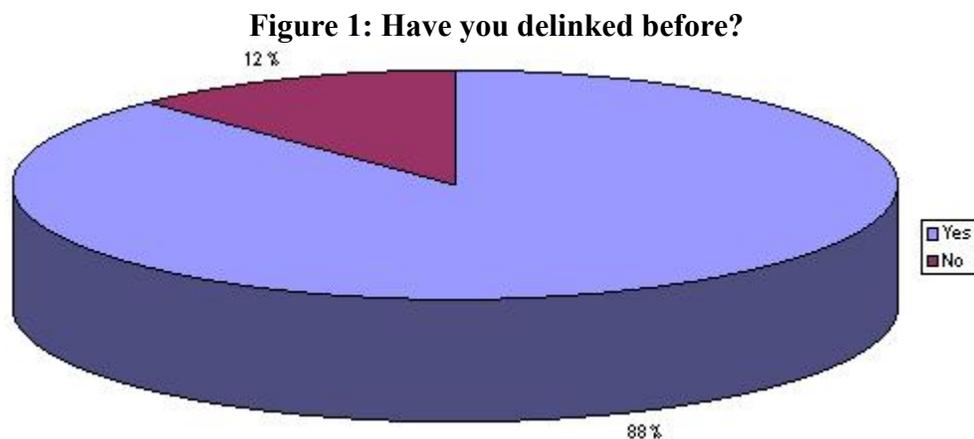
For this exploratory study, colleges and universities in Canada and the United States were chosen because of the availability of lists of their URLs. A list of URLs for American colleges and universities can be found at <http://www.clas.ufl.edu/au/> while a similar one for the Canadian ones can be found at http://www.aucc.ca/can_uni/our_universities/index_e.html

An online questionnaire, containing five simple questions, was created on SurveyMonkey.com. Two different collectors were also created to collect responses from American webmasters and Canadian webmasters. In the case of the Canadian web sites, each of the 92 sites was visited to obtain the e-mail address of the webmaster. A short e-mail message was then sent to these webmasters inviting them to fill the online questionnaire. In a few cases, the web site did not have an e-mail address for the webmaster but an online form for comments/suggestions. In such cases, the form was used to send the invitation to the webmaster. In the case of the American web sites, it was not feasible to visit them all as there are over 1500 of them. Given that the sites are arranged in alphabetical order by the name of the university/college, a systematic sample of 500 web sites was taken. Each of these web sites was then visited for the e-mail address or form to be used in inviting the webmaster to fill the online questionnaire.

Responses were collected over a one-month period with 25 responses (i.e. 27.2% response rate) received from Canadian webmasters and 98 (i.e. 19.6 % response rate) from the American webmasters.

Results

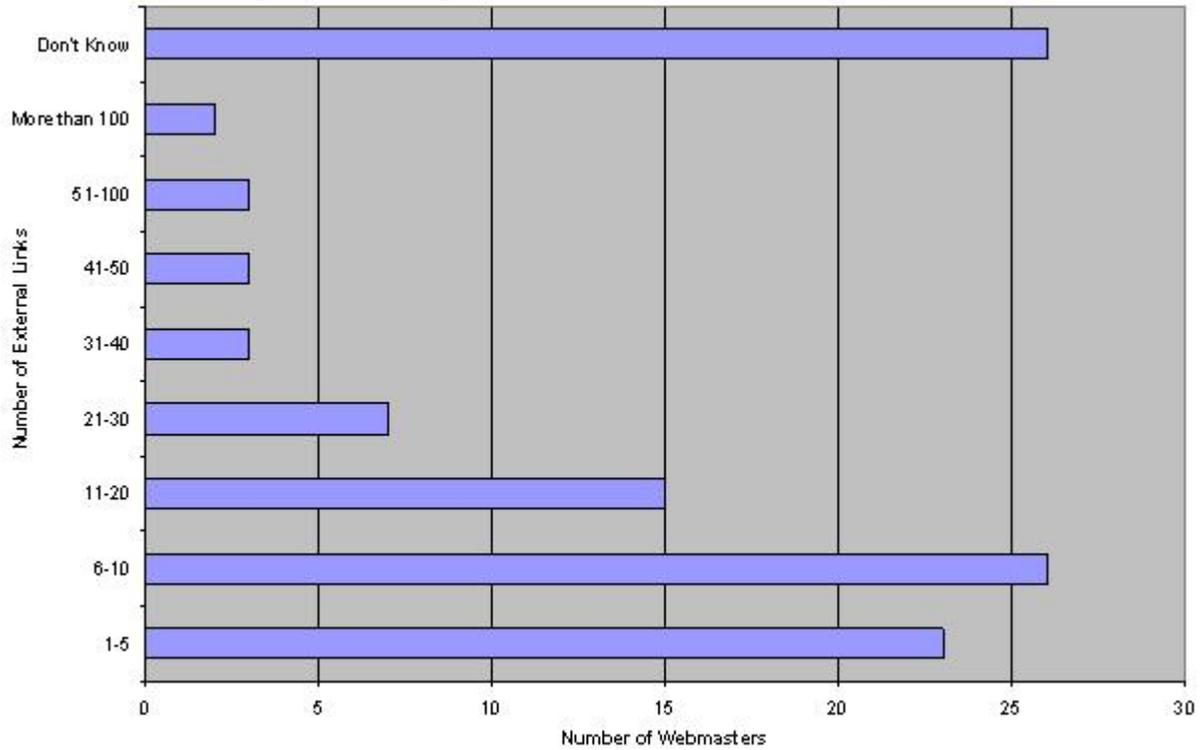
There was not much disparity between the responses between the American webmasters and their Canadian counterparts. Hence, a combined analysis is presented below.



Of the 123 respondents, about 88% have had the opportunity of removing external links from their web sites (see Figure 1). Of the few that were yet to remove external links, two indicated that they were new at their posts and were not sure of the past delinking practice of their sites. Another webmaster indicated that as a general policy, they do not link to external sites except for ticket purchasing sites; hence, removing external links is not a common practice for them. One other webmaster mentioned that they often correct broken links but he could not recall ever having to remove a link after it has been approved and posted. Of those who have delinked before, about 59% remove an average of 20 or less

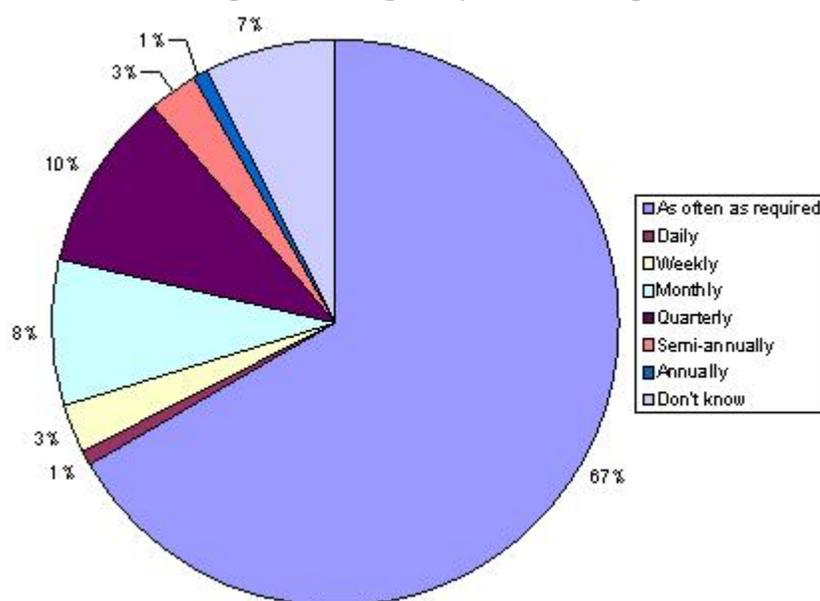
external links in a year, about 2 % remove more than 100 external links while about 24% could not give an estimate (see Figure 2).

Figure 2: Average number of links removed per year



In terms of how often delinking is done, almost two-third of the webmasters claimed to do it as often as required, especially whenever notified of a problem with an existing link or when re-designing a site (see Figure 3). According to one webmaster, "our procedure is very casual for removing external links. We usually are working on a page or content for other reasons when we discover external links that should be removed." About 26% of the webmasters claimed to delink on a regularly scheduled time basis, from daily to annually but monthly and quarterly are the two most popular time frequencies.

Figure 3: Frequency of delinking



As done in retraction studies, the reasons for delinking can be categorized into two, namely obvious (i.e. apparent), and unobvious. In the first category, and as expected, the number one reason given by webmasters for delinking is dead/broken/incorrect links (see Table 1). One pair of reasons has to do with the content of the site being linked to. About 71% of the webmasters indicated that they remove links if the content is no longer relevant, especially if the content has been changed. The other reason is if the content is no longer current. According to one webmaster, "delinking is important to keep the site fresh and up-to-date. Like collection development in a library - get rid of an out-dated book and replace it with a more current one" while another stated that "most external links are related to a news story or an event. Once the event has passed, the link is removed." Another pair of reasons has to do with the linking page. Links are automatically removed if the page linking is removed while links may also be removed if the content of the linking page is modified.

Table 1: Reasons for Delinking

Reason for Delinking	Number of Webmasters
Linked site is dead	99 (91.7%)
Content of linked site is no longer current	68 (63.0%)
Content of linked site is no longer relevant	71 (65.7%)
Political/philosophical difference with owner of linked site	15 (13.9%)
Linking page is removed	53 (49.1%)
Linking page is modified	55 (50.9%)
Others	8 (7.4%)

In the unobvious category, we have "political or philosophical difference with the owner of the linked site", and those labelled "others" in Table 1. These other reasons include: if the linked site's design or content is found to be of poor quality; if the content of the site linked to may be construed as an endorsement/partnership where none exists; if requested to remove the link by the public relations department of the owner of the linking site; or if requested to remove the link by the owner/webmaster of the linked site.

Conclusions

In this exploratory study, we have established that webmasters, especially those at colleges and universities in Canada and US, do delink on a regular basis, and reasons for delinking include dead links, site outdated, site is poorly designed, and site's content is of low quality. Given that most web search engines use the number of inlinks as one of the criteria in ranking search output, a site owner/webmaster should try to minimize the risk of the site being delinked from by other sites by ensuring that the site is well designed, rich in content, up-to-date, and if the site's (or any of its pages) address changes, users are re-routed from the old address to the new address.

It should be noted that the reasons given above for delinking are not specific to a particular web site but they pertain to the general practice by webmasters. To determine reasons for delinking from particular web sites (and possibly how frequently these sites are being delinked from), one may need to select a set of web sites, determine the external inlinks to each of them using a commercial search engine or a specially designed web crawler. Then, for a certain period of time, say 2 years, and at each interval, say monthly, the external inlinks to these sites will have to be determined again. For each site, the external inlinks for two successive periods will have to be compared to see the inlinks that have added or removed. In the case of inlinks removed, i.e. delinks, it would first have to be ascertained if they are dead. If they are not, then an e-mail will then have to be sent to the webmasters of these delinks to find out the reasons for delinking. The data collected from this kind of

study can also be used to develop a model for the growth of inlinks to a site over time. It would be interesting to see if the rate of delinking is significant enough to warrant including a delinking parameter in such a model.

Another issue related to delinking that is worth investigating in the future is the stability of authority sites. In a field of study, a site is considered an authority if it is pointed to by many other sites (Kleinberg, 1999). Based on this criterion, one could produce a ranked list of authority sites in a field. In the face of delinking, it would be interesting to study the stability of the ranks of authority sites in a field over time.

References

- Associated Press (2002). [Authors Guild: De-Link Amazon](http://www.wired.com/techbiz/media/news/2002/04/51676). Retrieved February 15, 2008, from <http://www.wired.com/techbiz/media/news/2002/04/51676>
- Bar-Ilan, J. (2004). A microscope link analysis of academic institutions within a country: The case of Israel. *Scientometrics*, 59(3), 391-403.
- Bjorneborn, L., & Ingwersen, P. (2001). Perspectives of webometrics. *Scientometrics*, 50(1), 65-82.
- Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30(1-7), 107-117.
- Budd, J.M., Sievert, M., Schultz, T.R., & Scoville, C. (1999). Effects of article retraction on citation practice in medicine. *Bulletin of the Medical Library Association*, 87(4), 437-443.
- Chu, H. (2005). Taxonomy of inlinked web entities: What does it imply for webometric research? *Library and Information Science Research*, 27(1), 8-27.
- Faba-Perez, C., Guerrero-Bote, V.P., & De Moya-Anegon, F. (2004). Methods of analysing web citations: A study of web-coupling in a closed environment. *Libri*, 54(1), 43-53.
- Garfield, E., & Welljams-Dorof, A. (1990). The impact of fraudulent research on the scientific literature - The Stephen E. Breuning case. *Journal of the American Medical Association*, 263(10), 1424-1426.
- Ingwersen, P. (1998). The calculation of web impact factors. *Journal of Documentation*, 54(2), 236-243.
- Kim, H.K. (2000). Motivations for hyperlinking in scholarly electronic articles: A qualitative study. *Journal of the American Society for Information Science and Technology*, 51(10), 887-899.
- Kleinberg, J.M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5), 604-632.
- 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.
- Kousha, K., & Thelwall, M. (2006). Motivations for URL citations to open access library and information science articles. *Scientometrics*, 68(3), 502-517.
- Larson, R.R. (1996). [Bibliometrics of the World Wide Web: An exploratory analysis of the intellectual structure of cyberspace](http://sherlock.berkeley.edu/asis96/asis96.html). *Proceedings of ASIS96*, pp. 71-78. Retrieved February 15, 2008, from <http://sherlock.berkeley.edu/asis96/asis96.html>
- Lohse, G.L., & Spiller, P. (2003). Internet retail store design: How the user influences traffic and sales. In: Steinfield, C., ed., *New directions in research on E-commerce*. West Lafayette, Indiana: Purdue University Press, 11-32.
- Nath, S.B., Marcus, S.C., & Druss, B.G. (2006). Retractions in the research literature: misconduct or mistakes? *Medical Journal of Australia*, 185(3), 152-154.
- Neale, A.V., Northrup, J., Dailey, R., Marks, E., & Abrams, J. (2007). Correction and use of biomedical literature affected by scientific misconduct. *Science and*

- Engineering Ethics*, 13(1), 5-24.
- Pfeifer, M.P., & Snodgrass, G.L. (1990). The continued use of retracted, invalid scientific literature. *Journal of the American Medical Association*, 263(10), 1420-1423.
 - Prime, C., Bassecoulard, E., & Zitt, M. (2002). Co-citations and co-sitations: A cautionary view on an analogy. *Scientometrics*, 54(2), 291-308.
 - Rousseau, R. (1997). [Sitations: An exploratory study](#). *Cybermetrics*, 1(1), paper 1. Retrieved February 15, 2008, from <http://www.cindoc.csic.es/cybermetrics/articles/v1i1p1.html>
 - Smith, A.G. (1999). A tale of two web spaces: Comparing sites using web impact factors. *Journal of Documentation*, 55 (5), 577-592.
 - Smith, A.G. (2004). [Web links as analogues of citations](#). *Information Research*, 9(4), paper 188. Retrieved February 15, 2008, from <http://InformationR.net/ir/9-4/paper188.html>
 - Tague-Sutcliffe, J. (1992). An introduction to informetrics. *Information Processing & Management*, 28(1), 1-3.
 - Thelwall, M. (2000). Web impact factors and search engine coverage. *Journal of Documentation*, 56(2), 107-122.
 - Thelwall, M. (2003). [What is this link doing here?](#) Beginning a fine-grained process of identifying reasons for academic hyperlink creation. *Information Research*, 8(3), paper 151. Retrieved February 15, 2008, from <http://informationr.net/ir/8-3/paper151.html>
 - Thelwall, M., & Wilkinson, D. (2004). Finding similar academic web sites with links, bibliometric couplings and colinks. *Information Processing & Management*, 40(3), 515-526.
 - Vaughan, L., Gao, Y., & Kipp, M. (2006). Why are hyperlinks to business websites created? A content analysis. *Scientometrics*, 67(2), 291-300.
 - Vaughan, L., & Hysen, K. (2002). Relationship between links to journal web sites and impact factors. *Aslib Proceedings*, 54(6), 356-361.
 - Vaughan, L., & Shaw, D. (2003). Bibliographic and web citations: What is the difference? *Journal of the American Society for Information Science and Technology*, 54(14), 1313-1322.
 - Whitley, W.P., Renne, D., & Hafner, A.W. (1994). The scientific community response to evidence of fraudulent publication - The Slutsky, Robert case. *Journal of the American Medical Association*, 272(2), 170-173.
 - Wilkinson, D., Harries, G., Thelwall, M., & Price, E. (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), 59-66.
 - Wolfram, D. (2000). Applications of informetrics to information research. *Informing Science*, 3(2), 77-82.

Bibliographic information of this paper for citing:

Ajiferuke, Isola (2008). "Delinking: An Exploratory Study." *Webology*, 5(1), Article 51.
Available at: <http://www.webology.org/2008/v5n1/a51.html>

Alert us when: [New articles cite this article](#)

Copyright © 2008, Isola Ajiferuke.