Bibliometrics and beyond: some thoughts on web-based citation analysis

Blaise Cronin

Journal of Information Science 2001; 27; 1
DOI: 10.1177/016555150102700101

Bibliometrics and beyond: some thoughts on web-based citation analysis

Blaise Cronin

Indiana University, USA

Received 5 September 2000
Revised 13 September 2000

Abstract.

The idea of a unified citation index to the literature of science was first outlined by Eugene Garfield [1] in 1955 in the journal *Science*. *Science Citation Index* has since established itself as the gold standard for scientific information retrieval. It has also become the database of choice for citation analysts and evaluative bibliometricians worldwide. As scientific publication moves to the web, and novel approaches to scholarly communication and peer review establish themselves, new methods of citation and link analysis will emerge to capture often liminal expressions of peer esteem, influence and approbation. The web thus affords bibliometricians rich opportunities to apply and adapt their techniques to new contexts and content: the age of 'bibliometric spectroscopy' [2] is dawning.

A retrospect

Bibliometricians count and measure things. Traditionally, they have concentrated their efforts on tracking highly visible and objective indicators of scholarly activity; most notably, publications and citations. Bibliometric techniques can, of course, be applied to many other inputs, outputs, processes and artifacts associated with the conduct of science (journals, acknowledgements, scientific manpower, federal funding patterns, rates of patenting, etc); in which case, either 'informetrics' or 'scientometrics' may be a more apposite descriptor. However, the models and methods used by these intersecting research communities, members from each of which belong to the recently established International Society for Scientometrics and Informetrics [46], have much in common and some of the subtle procedural distinctions may well blur as all three groups strive to develop context-sensitive methodologies and metrics appropriate to web environments.

Bibliometric analysis predates the development of *Science Citation Index* (*SCI*), but the advent of *SCI*, and specifically the availability of electronic access (online, CD-ROM and web-based) to the Institute for Scientific Information's (ISI) massive datasets, has had a catalytic effect on the popularity, scope and ambition of bibliometric research, both within and also well beyond the information science community. *SCI* grew out of a specialty index to the literature of genetics and was, as Garfield readily concedes, inspired by Shephard's legal citation index, which was created almost a century earlier [3]. Others, in fact, have argued that the true conceptual origins of citation indexing are to be found in fourteenth-century Hebrew literature [4]. For present purposes, a very brief (and necessarily impressionistic) history of citation indexing, including some derivatives and selected applications, is depicted in Fig. 1: in essence, the elements to the left of the dashed, vertical line inserted at approximately 1990. To the right of that line lie a number of current developments and proto-services (e.g. ResearchIndex, PubMed Central), which may, in due course, help to stimulate a bibliometric renaissance in the context of the web.

Originally, *SCI* was conceived as a practical tool for indexing and retrieving the literature of science, but it was not long before the full import and ramified significance of Garfield's invention began to be appreciated by scholars in a number of disciplines [5]. Subsequent

Correspondence to: Professor B. Cronin, School of Library and Information Science, Indiana University, Bloomington 47405, Indiana, USA. Tel: +1 812 855 2848. E-mail: bcronin@indiana.edu

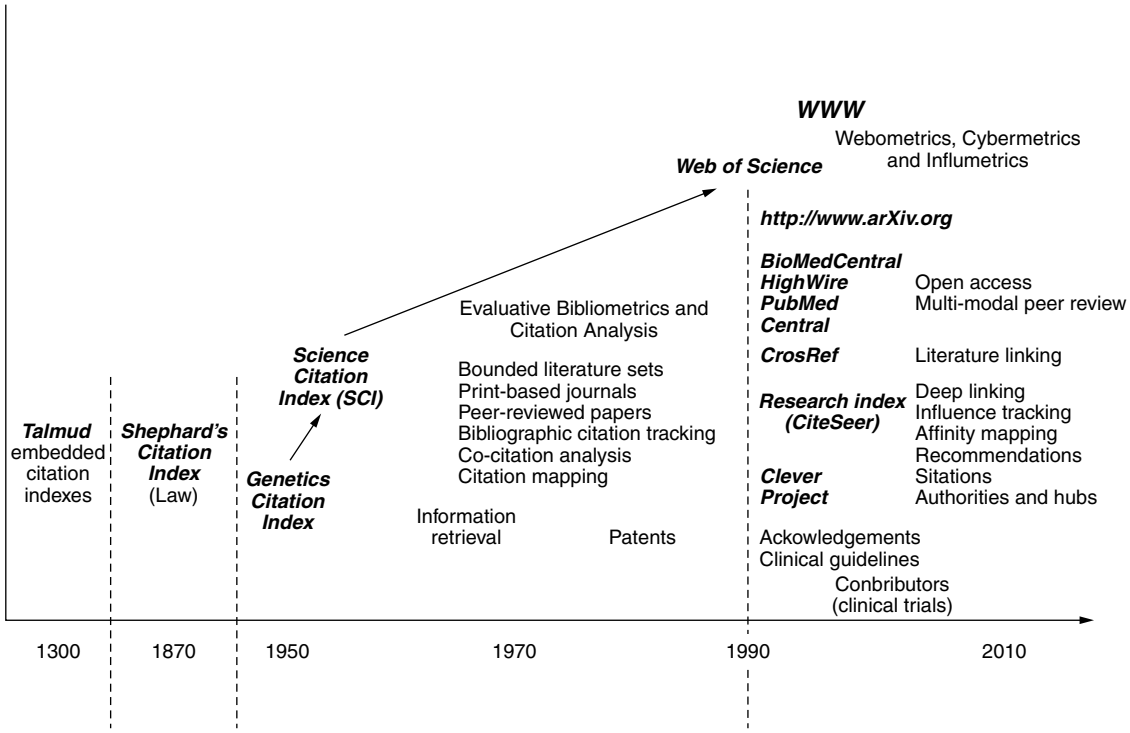

Fig. 1. The evolution of citation indexing.

development of sophisticated techniques for representing and mapping networks of scientific papers and authors (e.g. co-citation analysis) have taken us well beyond the principles and practice of information retrieval, the commercial *raison d'être* of *SCI*, and into the realms of sociometry, historiography and science policy. Over the course of the last few decades, Garfield's innovative retrieval system begat a slew of unintended and unforeseen applications and there is the prospect of much more to come.

Prospects

Citation analysis is an important piece of the bibliometric research pie; one that will become even more central with the growth of the web and for a very simple reason. The links (reference citations) provided routinely by authors in their reports and papers are a means of exposing the underlying socio-cognitive structure of science. However, links are also the defining feature of hypertext systems. As Larson [6] notes, the 'notion of citation is fundamental both to the scholarly enterprise and to hypertext networks where it provides the primary mechanism for connection and traversal of the information space (or "cyberspace")'. The principles of citation indexing find their echo in the dynamically reticulated structure of the web [7], hence the proliferation of neologisms, such as cybermetrics, netometrics, webometrics and influmetrics – the last of which was coined by Elisabeth Davenport to suggest diffuse and often imperceptible traces of scholarly influence – to capture the opportunities for measurement and evaluation afforded by the new environment. If citations can be tracked, counted and weighted, then why not the links connecting websites? After all, citations and 'sitations' are not merely similar phonetically, as Rousseau [8] has noted: appropriately enough, in *Cybermetrics,* a (pure) electronic journal. Highly linked sites are the web's equivalent of highly cited papers.

This thinking is axial to the work of IBM's Clever Project [9], the team members of which, in their efforts to map the web, explicitly acknowledge a methodological debt to Eugene Garfield and citation analysis. Specifically, they apply ISI's (journal) impact factor (see [47]) to the evaluation of websites in order to identify 'hubs' and 'authorities', the web's analogues of citation 'stars', nodes or centroids in classical bibliometric studies. Others have come to the same conclusion; the search engine Google (see [48]) uses link-based relevance weights to rank output. Going one step further, frequency of links, as with frequency of citation, can be

construed as a proxy for social trust, as has recently been acknowledged both within and outside the literature of information science [10, 11].

On the web, a diverse array of sites, people and objects can be linked to by others. This, potentially, facilitates the generation of multidimensional profiles of a scholar's or researcher's presence within (and across) particular communities of practice. As has been noted elsewhere [12], the 'range of genres of invocation made possible by the Web should help give substance to modes of influence which have historically been backgrounded in narratives of science, or, at best, picked up, parenthetically, in biographies of the great and the good'. In other words, more of the world's 'silent scientists', to use Meadows' [13] label for those who are not particularly visible in terms of their publication or citation records, may finally receive due recognition for some of their unseen and unsung contributions. In addition to citations, we can track other forms of invocation, which, individually or conjointly, may provide more finely textured accounts of an individual's (or research team's) influence on peer group thinking, both locally and globally.

New publishing environments

The significance of the web from a bibliometric perspective goes well beyond enhanced opportunities for citation and link analysis. The web has challenged, and may revolutionize, many of the assumptions that have underpinned the established scholarly communication system. Radical proposals for open access (self-publishing, self-posting and self-archiving in whatever configuration) and open peer review (the bypassing of traditional refereeing approaches) have been put forward by a growing band of scholars and researchers, perhaps most cogently and compellingly in a blizzard of papers and presentations over the past few years by the often co-cited trinity of Stevan Harnad, Andrew Odlyzko and Paul Ginsparg, the last of whom also happens to be the founder of the enormously successful physics e-pre-print archive at Los Alamos [49], a model which is being copied, albeit selectively, in a number of other domains. Support for their buccaneering ideas comes not only from significant sections of the grass-roots scientific community [14], but occasionally from the apex of the scientific establishment.

Harold Varmus, until recently Director of the National Institutes of Health (NIH), the federal agency responsible for funding $16 billion of biomedical research annually in the USA, was instrumental in conceiving and launching the (continuously evolving) proposal to establish PubMed Central (*né* E-biomed). The aim, in short, is to create a web-based repository of biomedical research literature hosted by the NIH, to which global access will be provided free (see [50]). Commercial publishers are, naturally, less than thrilled by proposals for free and open access to the literature of biomedicine and cognate fields. The recent creation of CrossRef by a consortium of some 30 commercial publishers – it uses the digital object identifier (DOI) to link reference citations to full-text, online content held by different publishers (see [51]) – may be construed, in part, as a response to unorthodox (and essentially public sector-inspired) proposals for open access and archiving. CrossRef, which describes itself as 'the comprehensive source for linking journal articles', also provides further evidence of the fact that hypertext and citation indexing are a marriage made in heaven, though ISI may well view this pan-publisher initiative as a potential strategic threat to its *de facto* monopoly on the provision of large-scale, longitudinal citation data. Another pertinent – and, again, from ISI's commercial perspective, potentially threatening – development is the emergence of autonomous citation indexing as embodied in ReferenceIndex (formerly CiteSeer: see the right-hand side of Fig. 1), a value-adding system, developed at the NEC Research Institute in Princeton, that automatically extracts citations *and the context* in which the citations are made in the body of the citing paper [15]. This allows the reader to see exactly how the citation relates to the surrounding text and to better understand its instrumental and symbolic significance.

What is unclear in all of the ongoing experimentation and attendant speculation is the extent to which established peer review practices may be subverted by open peer review or, indeed, the extent to which different scientific communities and sub-fields may wish to adopt a tiered model of peer review, ranging from, say, full-bodied, double-blind refereeing through 'peer review lite' to zero peer review. It does, however, seem likely that the present monolithic publishing and peer review system will become rather more flexible and pluralistic in character, as new publishing, posting, critiquing and archiving behaviours establish themselves. Assuming that a more diverse publishing environment gradually emerges, bibliometricians will have a much broader array of objects and artifacts to feed into their accounts and analyses – both quantitative and qualitative – of scholars' communicative practices. This, inevitably, will turn the spotlight on a number of important issues relating to the provenance and durability (both intellectual and physical) of digital outputs.

Heterogeneity of output

Traditionally, bibliometricians have dealt with an inherently stable environment – print-based publishing – and largely standard artifacts, outputs and units of analysis (e.g. scholarly journals, peer-reviewed articles, citations). The new publishing environment, by way of contrast, promises heterogeneity of output and therewith greatly enhanced scope for, *inter alia*, quantifying multiple (and multimedia) outputs, tracing socio-cognitive linkages and harvesting online commentary/glosses on the works and ideas of professional scientists and scholars. What, then, are the probable units of analysis associated with web-based publication and communication fora? What, in fact, will be tracked, counted, weighted and evaluated? In addition to the traditional journal article, we have electronic pre-prints, self-posted/self-archived documents, dynamically revised working papers, and multiple forms of 'scholarly skywriting', to use one of Harnad's [16] colourful metaphors. How, for instance, should we track, log and assess such 'sensitive measures of "air time" and flight rate for new ideas and findings'? [16]. We also need to know how to identify, locate and access these non-traditional (and, for now at least, often evanescent and ephemeral) outputs. Will common standards for identification, formatting and labelling emerge, as well as for assuring the stability and persistence of less significant digital objects (e.g. online syllabi, courseware, threaded online chat, digital drafts, citations) over time?

In the present system, scholarly articles 'belong to' particular journals, which, in the main, have stable existences and regular publication cycles. Libraries with more or less coherent policies for collection development and management provide persistent access to serials. In the new world order, bibliometricians and others will presumably want to know whether the objects of their scrutiny are part of a journal, host service, depository, discussion forum, website or electronic archive. They will also need to have some shared sense of the quality, authority and integrity of what is being mapped and measured. For instance, a scientific report hosted by the NIH's PubMed Central has, *prima facie*, higher credibility and cognitive authority than an unvetted opinion piece hosted on this author's personal server. All outputs are not created equal and all citations are not equivalent. The warranting of claims in open publishing environments will most certainly be a more complex and nuanced matter than has been the case in the pre-digital communication system of science.

Pedigree and persistence

In a pluralistic and promiscuous publishing environment, the pedigree, provenance and persistence of documents, as well as their hosts, will be issues of consequence. Quality, in the present system, is signalled in a variety of ways. For example, major abstracting and indexing services, such as Biosis, Medline or Chemical Abstracts, may cover a journal, or the published/cited work may acknowledge funding from a research council, foundation or other ostensibly reputable agency, which, typically, implies some degree of adjudication and also supports a reasonable presumption of quality. Our trust in texts is a function of the extent to which they have been screened or subjected to pre-publication peer review in some form. The idea of open peer review on the web, or differential peer review, is gathering momentum, though concerns that the positive features of the present system (e.g. filtering), its acknowledged flaws nothwithstanding, might be undermined, remain widespread.

As new technologies emerge and novel modalities of posting, hosting, publishing and filtering are developed, we can expect different disciplines and sub-fields to deploy quite different approaches and techniques. In addition, the rates of adoption of new models and behaviours will vary [17]. The material practice of science varies greatly from one domain to the next and assumptions of common values and comparable metabolic rates are ill-founded. The socio-cognitive structure of disciplines, and also their reward systems, can (and do) differ greatly, which helps to explain why, for instance, e-pre-print exchange is so readily accepted within the rarefied world of high energy physics but is still regarded with suspicion by some other scientific communities [18, 19]. Universalistic approaches to communication, publication, filtering and evaluation should certainly not be assumed: rather, pluralism, 'plasticity, and adaptivity will be the hallmarks of the new world order' [20].

Validity and reliability issues

Critics of citation analysis (see, for example, [21]) have long challenged the assumption that citations can be used as valid indicators of quality, utility or even impact, despite a compelling body of research findings to support the core contention (see, for example, [22, 23]). They have also identified reliability-related issues arising from the (perceived) selectivity and limitedness (e.g. in terms of language and geography) of ISI's

coverage, though autonomous citation indexing [15] may be a harbinger of unbounded and reliable citation indexing. However, even if open hypertext systems (as originally conceived by Cameron [24]) allow us to get around the limitations of ISI's coverage, there is still the underlying problem of search engine reliability. Not only is there considerable variation between search engine retrieval performances, but the same search engine will also generate different output for the same search at different times [12, 25, 26]. For this reason alone, extreme caution is advised in using web-derived, proto-indicators of scholarly or communicative salience for evaluative purposes, most especially at the level of the individual [12, 27].

Undoubtedly, though, construct validity issues will continue to surface, as new forms of web-based invocation are factored into bibliometric evaluations and sociometric narratives of scientific communication. Sceptics, as with the ever-vocal doubting Thomases of classical citation analysis [28], will, quite rightly, want to know what various kinds of 'sitations' and 'hits' signify [29] and how we should go about distinguishing the substantive from the seriously trivial in open publishing environments. Adler [30, 31] has coined the phrase 'Slashdot Effect' to describe the spontaneous hit rate on a web server following the posting of a news story to a high-traffic site. Such surge-activating papers, postings or provocations may often prove to have little, if any, enduring academic, or other, merit. In web advertising terminology, it is the difference between click-through and look-to-buy metrics: the former record the number of visitors referred to a site by an online banner advertisement, while the latter compares banner impressions with subsequent sales transaction data [32]. Furthermore, it has been found that website traffic estimates produced by firms such as Media Metrix and Nielsen NetRatings seem to significantly influence a company's stock market valuation [33]. How-ever, the velocity and intensity of much web-based communication can easily create a false sense of what matters within a given community of scholars. The lessons are clear and the challenge for bibliometricians in the new environment will be to understand that a Gadarene rush does not necessarily equate with a posting's scholarly or utilitarian significance.

Concluding remarks

Citation indexing, as Atkins [34] has observed, 'is a prime example of a concept before its time'. With the creation of the web, its time has unquestionably come. The application, and extension, of existing bibliometric techniques to open publishing environments should, ultimately, result in the development of a portfolio of metrics for better capturing the totality of ways in which cognitive influence is exercised and exhibited within and across specialty groups. One of the limitations of the ISI's databases is that they are restricted to a relatively small, albeit high-quality, sub-set of the universe of scholarly journals. This means that citations to a scholar's work that appear in journals (not to mention monographs [35]) excluded from ISI's coverage are, to all intents and purposes, lost. Additionally, journals not included in ISI's coverage are less visible than they might otherwise be: an understandably sore point for some peripheral nations, especially since many of the papers cited by the journals covered by ISI's indexes are actually published in journals which lie beyond the citational pale [36]. However, as the developers of ResearchIndex have shown in their prototype [15], there is a possibility that a new generation of citation indexing tools may emerge, i.e. one that is not subject to the kinds of commercial and technical constraints that have necessarily inhibited the development of ISI's citation indexing products.

The hypertextual character of the web means that the principles of citation indexing can be applied much more widely than at present. Web-based retrieval systems will allow us to go beyond traditional citations and track acknowledgements, diffuse contributions and related input measures: a fact quickly grasped by, among others, the Wellcome Trust [37]. It is also worth noting in this context that Drummond Rennie (Deputy Editor of *Journal of the American Medical Association*) and colleagues in the medical community have proposed that the concept of contributorship should replace that of authorship. Given the well-documented problems associated with fraud and honorific authorship in large-scale clinical trials, in particular, there is a growing movement to itemize individuals' specific contributions to a project and also to identify one or more members of the collaboration who would act as the guarantors of the resulting report or journal article [38, 39]. Broadly conceived, web-based citation

indexing and analysis should ensure that the sometimes overlooked inputs and influences of technicians, mentors, trusted assessors and sundry collaborators [40, 41] could more easily be factored into the recognition and assessment calculus. That said, great care will be needed in order to ensure that the symbolic significance of such signs – these 'pellets of peer recognition', to appropriate Merton's [5] euphonious phrase – is not misinterpreted and subsequently misused [42].

Citation context analysis and deep linking (pointing to a precise spot on a website) will allow us to explore the contexts in which invocations occur and thereby develop a more reliable sense of why an individual warrants mention or how an individual's work and ideas are perceived and received by his or her peer community. On the web, scholars do more than publish, or post, their working papers and finished articles: they 'seed ideas, discuss issues and debate positions, in ways which, occasionally deviate from, and challenge, established norms' [12]. Furthermore, they recommend their own work, and the work of selected others, to their peers. This, of course, is what we do (most of the time) when we cite others. Citation indexes may thus be viewed as a prototypical species of recommender system (see, for example, [43]). In the near future, web-based 'bibliometric spectroscopy' [2] will allow us to map with greater fidelity the matrix of informal endorsements and recommendations which lie behind the formal rating systems of science: 'social indicator mining', if you will. Concrete applications might include co-citation analysis of hyperlinks to reveal the organization of online communities and their knowledge bases [44]. It is equally conceivable that next-generation search engines and bibliometric tools will allow us to observe 'science in action' [45]; in the process, enabling us to detect early signs of emerging trends and also develop a better sense of those scientists and researchers who have what might be termed 'street cred', as opposed to the established coin of the realm: eminence.

Acknowledgements

This paper grew out of an invited presentation at a conference, *Freedom of Information: The Impact of Open Access on Biomedical Science*, held at the New York Academy of Medicine, New York, 6–7 July, 2000 (for details, see [52]). I am grateful to Helen Atkins of ISI for helpful suggestions and to both referees for their comments.

References

[1] E. Garfield, Citation indexes for science: a new dimension in documentation through association of ideas, *Science* 122 (1955) 108–111.

[2] A.F.J. van Raan, The Pandora's box of citation analysis: measuring scientific excellence, the last evil? In: B. Cronin and H.B. Atkins (eds), *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield* (ASIS Monograph Series) (Information Today Inc, Medford, NJ, 2000), pp. 301–319.

[3] P. Wouters, *The Citation Culture* (PhD Thesis) (University of Amsterdam, 1999).

[4] B.H. Weinberg, The earliest Hebrew citation indexes, *Journal of the American Society for Information Science* 48(4) (1997) 318–330.

[5] R.K. Merton, On the Garfield input to the sociology of science: a retrospective collage. In: B. Cronin and H.B. Atkins (eds), *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield* (ASIS Monograph Series) (Information Today Inc, Medford, NJ, 2000), pp. 435–448.

[6] R.R. Larson, Bibliometrics of the World Wide Web: an exploratory analysis of the intellectual structure of cyberspace. In: S. Hardin (ed.), *Global Complexity: Information, Chaos and Control. Proceedings of the 59th ASIS Annual Meeting, Baltimore, MD, October 21–24, 1996* (Information Today Inc, Medford, NJ, and ASIS, 1996).

[7] H. J. Kim, Motivations for hyperlinking in scholarly electronic articles: a qualitative study, *Journal of the American Society for Information Science* 51(10) (2000) 887–899.

[8] R. Rousseau, Sitations: an exploratory study, *Cybermetrics* 1(1) (1997). Available at: http:// www.cindoc.csic.es/cybermetrics/vol1iss1.html

[9] Clever Project, Hypersearching the Web, *Scientific American* (June 1999) 54–60.

[10] E. Schlossberg, A question of trust, *Brill's Content* 2(2) (1999) 68–70.

[11] E. Davenport and B. Cronin, The citation network as a prototype for representing trust in virtual environments. In: B. Cronin and H.B. Atkins (eds), *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield* (ASIS Monograph Series) (Information Today Inc, Medford, NJ, 2000), pp. 517–534.

[12] B. Cronin, H.W. Snyder, H. Rosenbaum, A. Martinson and E. Callahan, Invoked on the Web, *Journal of the American Society for Information Science* 49(14) (1998) 1319–1328.

[13] A.J. Meadows, *Communication in Science* (Butterworths, London, 1974).

[14] B.P. Markovitz, Biomedicine's electronic publishing paradigm shift: copyright policy and PubMed Central, *Journal of the American Medical Informatics Association* 7 (2000) 222–229.

[15] S. Lawrence, C.L. Giles and K. Bollacker, Digital libraries and autonomous citation indexing, *IEEE Computer* 32(6) (1999) 67–71.
[16] S. Harnad, Interactive publication: extending the American Physical Society's discipline-specific model for electronic publishing, *Serials Review* 18(1–2) (1992) 58–61.
[17] S.P. Harter and T.K. Park, Impact of prior electronic publication on manuscript consideration, *Journal of the American Society for Information Science* 51(10) (2000) 940–948.
[18] R. Kling and G. McKim, Not just a matter of time: field differences in the shaping of electronic media in supporting scientific communication, *Journal of the American Society for Information Science* (2000) [in press].
[19] B. Cronin, Hyperauthorship: a postmodern peculiarity or evidence of a structural shift in scholarly publishing practices? [Submitted for publication.]
[20] B. Cronin, Will e-publishing save academic research? *The Chronicle of Higher Education* (15 October 1999) A25.
[21] M.H. MacRoberts and B.R. MacRoberts, Problems of citation analysis: a critical review, *Journal of the American Society for Information Science* 40(5) (1989) 342–349.
[22] S. Baldi, Normative versus social constructivist processes in the allocation of citations: a network analytic model, *American Sociological Review* 63 (1998) 829–846.
[23] R.R. Braam, *Mapping of Science: Foci of Intellectual Interest in Scientific Literature* (DSWO Press, University of Leiden, 1991).
[24] R.D. Cameron, *A Universal Citation Database as a Catalyst for Reform in Scholarly Communication* (1997). Available at: http://elib.cs.sfu.ca/project/papers/citebase/ citebase.html
[25] H. Snyder and H. Rosenbaum, Can search engines be used as tools for web-link analysis? A critical view, *Journal of Documentation* 55(4) (1999) 375–384.
[26] R. Rousseau, Daily time series of common single word searches in AltaVista and Northern Light, *Cybermetrics* 2/3(1) (1998/99). Available at: http://www.cindoc.csic.es/cybermetrics/ articles/v2i1p2.html
[27] B. Cronin, The Warholian moment and other proto-indicators of scholarly salience, *Journal of the American Society for Information Science* 50(10) (1999) 953–955.
[28] G. Taubes, Measure for measure in science, *Science* 260 (14 May 1993) 884–886.
[29] P. Ingwersen, The calculation of web impact factors, *Journal of Documentation* 54(2) (1998) 236–243.
[30] S. Adler, *The Slashdot Effect: An Analysis of Three Internet Publications* (1999). Available at: http://ssadler.phy.bnl.gov/adler/SDE/SlashDotEffect.htm
[31] S. Adler, *Addendum to the Slashdot Effect Internet Paper* (1999). Available at: http://ssadler.phy.bnl.gov/adler/SDE/SlashDotEffectAddendum.html
[32] E. Schonberg *et al.*, Measuring success, *Communications of the ACM* 43(8) (2000) 53–57.
[33] M.J. Thompson, S. Lawrence, D. Lake and M.A. Mowrey, Trading on ratings, *The Industry Standard* (24 July 2000) 171, 173–175, 178–179, 182–183, 186, 188, 190, 193.
[34] H. Atkins, The ISI *Web of Science* – links and electronic journals: how links work today in the *Web of Science*, and challenges posed by electronic journals, *D-Lib Magazine* (September 1999). Available at: http://www.dlib.org/dlib/september99/atkins/09atkins/html
[35] B. Cronin, H. Snyder and H. Atkins, Comparative citation rankings of authors in monographic and journal literature: a study of sociology, *Journal of Documentation* 53(3) (1997) 263–273.
[36] S. Maricic, Mainstream-periphery science communication, *Nature* (28 January 1999). Available at: http://helix.nature.com/wcs/e01.html
[37] G. Lewison, Citations as a means to evaluate biomedical research. In: B. Cronin and H.B. Atkins (eds), *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield* (ASIS Monograph Series) (Information Today Inc, Medford, NJ, 2000), pp. 361–372.
[38] D. Rennie, V. Yank and L. Emanuel, When authorship fails: a proposal to make contributors accountable, *Journal of the American Medical Association* 287(7) (1997) 579–585.
[39] D. Rennie and V. Yank, If authors became contributors, everyone would gain, especially the reader, *American Journal of Public Health* 88(5) (1998) 828–830.
[40] N.C. Mullins, *Theories and Theory Groups in Contemporary American Sociology* (Harper and Row, London, 1973).
[41] B. Cronin, *The Scholar's Courtesy: The Role of Acknowledgement in the Primary Communication Process* (Taylor Graham, London, 1995).
[42] B. Cronin, Semiotics and evaluative bibliometrics, *Journal of Documentation* 56(4) (2000) 440–453.
[43] L. Terveen, W. Hill, B. Amento, D. McDonald and J. Creter, PHOAKS: a system for sharing recommendations, *Communications of the ACM* 40(3) (1997) 59–62.
[44] J. Pitkow and R. Pirolli, Life, death, and lawfulness on the electronic frontier. In: *Proceedings CHI 97, Conference on Human Factors in Computing Systems, 22–27 March 1997, Atlanta, Georgia, USA*. Available at: http://www.acm.org/sigchi/chi97/proceedings/paper/ jp-www.htm
[45] M.J. Bossy, The last of the litter: 'Netometrics', *Solaris* 2 (1995). Available at: http://www.info.unicaen.fr/bnum/jelec/Solaris/d02/2bossy.html
[46] Available at: http://crrm.u-3mrs.fr/issi/issi-home.html
[47] Available at: http://www.isinet.com/isi/hot/essays/7
[48] Available at: http://www.google.com/promote.html #search
[49] Available at: http://www.arXiv.org
[50] Available at: http://www.pubmedcentral.nih.gov/
[51] Available at: http://www.crossref.org/
[52] Available at: http://www.biomedcentral.com/info/conference.asp