

Short paper for ASIST-AoIR workshop 'The web as a mirror of scientific and technical achievements: issues in access and measurement' in Brighton, **19 September 2004**

## **Qualitative Aspects of Quantitative Measurements in the Age of Cyberscience**

**Michael Nentwich**

[mnent@oeaw.ac.at](mailto:mnent@oeaw.ac.at)

Institute of Technology Assessment,  
Austrian Academy of Sciences, Vienna  
<http://www.oeaw.ac.at/ita/>

### **Abstract**

Cyberscience will be different from traditional science. For two decades now, the scholarly community has witnessed a considerable increase in the use of information and communication technologies (ICT). Unlike 'traditional' science that does without networked computers, the notion of 'cyberscience' captures the use of these ICT-based applications and services for scientific purposes. It can be shown that ICT use impacts on the basic parameters of how academia is organised, of how it functions, and of what it produces (Nentwich 2003). We are in the middle of a forceful development. Cyberscience is already taking place, but will develop its full shape and potentials only at a later stage. The new media have only just begun to play a central role in a large array of scholarly activities, and in regard to the institutional setting. Not only academic communication in the narrow sense, but also the distribution of knowledge and, most importantly, even knowledge production are affected. Hence, the impact of ICT can hardly be underrated.

In the first part of my presentation, I give an overview on my book on 'Cyberscience – Research in the Age of the Internet', which describes and analyses the use of ICT in the academic world, explains the status quo based on an analytical model, draws a differentiated picture of probable future developments, assesses the impact of ICT on various aspects of academic activity and on the substance of research, and discusses the implications for research policy and the steering mechanisms within the scholarly organisations.

In the second part, I focus on the prospects of new methods of ex-post quality control in academia. Rating, scoring, commenting, access counts, citation counts and use tracking need to be looked at not only from a technical perspective of feasibility. New aspects such as privacy have to be taken into account as well as known issues in new clothes such as the types and significance of citations.

## **I. Short Overview on “Cyberscience – Research in the Age of the Internet”**

The book on ‘Cyberscience – Research in the Age of the Internet’ came out in November last year with the Austrian Academy of Sciences Press in Vienna. It is also available online as an E-book – a few chapters even for free. It comes with a large annotated and searchable link database. The book is the main outcome of a three years project, carried out at the Institute of Technology Assessment in Vienna and at the Max Planck Institute for the Study of Societies in Cologne.

On its 600 pages I describe and analyse the use of ICT in the academic world in an encompassing way. The book tries to explain the status quo based on my own analytical model. I draw a differentiated picture of probable future developments, assess the impact of ICT on various aspects of academic activity and on the substance of research, and discuss the implications for research policy and the steering mechanisms within the scholarly organisations.

The plan of the book is as follows:

In a first step, I collected data on the status quo. E-mail is used practically universal and online databases are spread very widely. While E-mail discussion lists are partly considered very important, Internet video-conferencing is, so far, the rare exception. Virtual research institute exist only to some extent, but there is a trend to more. Electronic journals, full text databases for digital working papers, as well as digital libraries are gaining ground. Altogether, I elaborated and compared 13 case studies of academic disciplines. Contrary to expectations, not only the natural sciences use increasingly these new technologies. Also several fields in the humanities, social sciences and cultural studies, such as parts of history and papyrology, are well advanced in this respect.

In a second step, I analysed in-depth what kinds of impact the use of new media already has today, and which effects are to be expected in the future. In cyberscience, spatial distances (e.g. with regard to libraries) become less important. Peripheral research distant to the centres will be easier to some extent. In the field of academic publishing, in particular, important changes are looming. In academia, electronic publishing will replace print and paper in the not so distant future. New procedures of quality control are in the making and will complement the existing ones. New formats of knowledge representation (hypertext, multimedia) are likely to be seen soon next to the classical scientific article or book. In addition, the much cheaper electronic publishing opens up the opportunity that academia gets eventually rid of the financial burdens that are laid upon it by the present commercialised publishing system. As regards copyright law, there is urgent need for action in order to guarantee research the necessary access to all important information. Finally, the study found that, potentially, there will be even indirect impact on the substance of research. The electronic media trigger changes in methodologies, work modes and as regards the representation of knowledge. These changes may have an impact on knowledge production.

For more information on the project and the book please refer to this website:

<http://www.oeaw.ac.at/ita/cyberscience.htm>

## II. Assessment of cyberscience methods of ex-post quality control<sup>1</sup>

In the traditional paper-based academic publishing environment, only ex ante quality control is feasible. If something turns out to be erroneous or to need qualification, it nevertheless stays part of the published “body of knowledge”. There is no way to delete ex post a text printed and distributed around the world, nor would it possible to tag the contribution saying that the content “should be handled with care”. Cyberscience, instead, provides for a variety of ex post mechanisms that could do exactly this.

As a first alternative, I should mention that already today (and even more so in the future), most researchers access publications not directly through browsing the shelves of a library, but by scanning the results of an online query in a bibliographic database. One simple, but dangerous way of deflecting from false results—no matter how this might have been established—would be to remove the item from the “first layer” of the database. The effect of this would be that the item would not be found any more. Other alternatives possible in the digital environment leave the publication and the bibliographic entry untouched, but let readers (as well as the author) attach comments of various sorts. Subsequent readers may use this added information for their own assessment of the content. In the following, I shall discuss three such novel mechanisms: annotation (1), rating (2) and use tracking (3).

### **(1) Online (ex post) commenting/annotation**

Quality control might be provided after publication by reader comments (or annotations) that the new technology could “attach” to any given article. This may be organised for self-published manuscripts as well as for articles formally published in E-journals (whatever the original quality check). The necessary technology is already developed and various experiments are under way. In the Semantic Web initiative, the WWW Consortium has taken up this idea in its ambitious Annotea project. The idea is that positive comments would probably make an article more significant, while negative comments (or no comments) would probably encourage its marginalisation. LaPorte et al. (1995) exemplify this with the plan of a “global health information server” where papers would get a public comment card: “If papers are poor, then the scientific community will most certainly indicate that they are poor; this is the nature of science, and this is the nature of the internet” (for law, see Hibbitts 1996; Tomlins 1998). Each review would rate an article without (as in the current system) forcing its (perhaps premature) suppression due to a negative verdict of the reviewers. In such a system, each reader can have the critical warrant that has been the prerogative of the traditional referee (Davenport/Cronin 1990).

The comments are stored together with (or at least directly linked to) the paper which becomes a “living document” (Odlyzko 1994). This could be combined with a system allowing the author to submit, from time to time, a new version of the paper including or taking due account of the comments.

The subversive potential of ex post commenting challenges how the questions of “what is scholarship and who is a scholar” (Nathenson 2001, 3) are decided today. Whether the quality of the collective assessment—which certainly includes evaluations of people with insufficient knowledge—meets the quality of appointed referees is an open question. It reminds us of the debate about democracy and whether, for example, a plebiscite can be a valid alternative to expert decisions. This is not only an ideological question. What we can learn from this parallel discussion regarding the political sphere is, inter alia, that minimum thresholds may play a role. Similar to political votes which may not be valid because too few people

---

<sup>1</sup> This part draws on Nentwich (2003), chapter 6, and Nentwich (2004).

participated in the voting, one could think of designing ex-post quality control systems that do show comments only as soon as a minimum number of people have registered their comments. This is, however, a very tricky issue and needs in-depth pondering and testing. A related issue to consider is that commenting may lead to what may be called “mainstreaming”. While at the same time enabling dissenting opinions to get published at all, “fishing” for positive comments and ratings may tend to favour less controversial contributions. Again, the effect is possible, but difficult to be evaluated in advance without practical experience.

## **(2) Rating / scoring**

Commenting is but one form of reacting to published manuscripts. It has two “disadvantages”: first, commentators have to write new text which takes time and needs more in-depth reflection; second, subsequent readers, too, need time for orientation as they have to browse through all comments with a view to making sense of all of them before being in a position to decide whether they should go on reading the main article. Either in combination with commenting or independently, “rating” (scoring) has been proposed to solve these two problems. Rating is giving marks by readers (Alton-Scheidl et al. 1997). The marks are collected and means are computed. The current results are offered to the reader (similar to online bookstores such as Amazon). Some argue quite radically that sophisticated, democratically legitimised rating systems could deprive the “traditional evaluation elites” of their power (Rost 1998). For journal articles, it is conceivable to combine a rough screening procedure to exclude “libellous, salacious, or otherwise unsuitable material” (Varmus/et al. 1999) with a commenting and rating procedure.

The problem with rating is that it is difficult to control for participation. If everyone, including the author, may rate, even more than once, there is a big chance that we would end up with a strong bias and the result would be highly doubtful, if not unusable. In a closed system with all possibly participating researchers known and with strict access control (everyone having a user-id and password), it would be technically feasible to implement a semi-anonymous system. Such a system would allow checking that each member of the closed group (except the author) has exactly one voice without the individual ratings being de-anonymised. However, it would probably be a major task to administer the groups’ membership—a task for scholarly associations. Another problem would be how to cope with controversial results. If a manuscript receives only positive or only negative ratings, a system computing averages would be helpful. But what about a paper that splits the community and gets very positive and very negative scores. The resulting medium score would hide more than it would reveal if the standard deviation (variance) would not be indicated, too. Another solution to this problem would be a more differentiated scoring system, for instance, a multi-dimensional rating scheme that is not only one overall score. Most advocates of scoring systems in academia, however, tend to see it as an add-on to other systems controlling for quality. The same is also true for the next innovation.

## **(3) Access and citation counts, use tracking**

The quality of a publication can be and is already measured by its use, as well. In the paper world, use can solely be computed by counting sold copies of a journal or book. This is, however, very imprecise, as copies of journals or edited books contain several articles and as many journals go to libraries which have an indefinite number of users. Furthermore, it is indirect, as the sole fact that a copy is within reach is only a very rough indication that the article is actually read and used. Today, counting citations of a particular article in other publications is a de-facto standard and specialist organisations such as ISI compute impact factors. While citations are already a better indication of the actual use of a publication, this

method is still not a perfect one as the reason for citing an article can be quite diverse (for an overview on the critique of present scientometrics, see e.g. Fröhlich 1999; Wouters 1999).

The advent of E-publishing has brought about new perspectives for scientometrics:

- (a) If all publications included in the citation analysis are available online, and perhaps even cross-linked, it is much easier and more efficient to do the citation calculations. In this context, Cameron (1997) proposed a universal, Internet-based, bibliographic and citation database (“citebase”) that would “link every scholarly work ever written—no matter how published—to every work that it cites and every work that cites it”.
- (b) A hypertext environment with standardised meaningful (citation) links would be in an even better position to make the bibliometric enterprise useful (Davenport/Cronin 1990, 187f.).
- (c) Unlike in the paper world, there may be automatically generated figures available telling us how many people have already accessed and/or downloaded a particular digital document.
- (d) Finally, use tracking is more than just counting access; it is tracing and computing the degree to which different publications are used and in what ways (Atkinson 2000, 67). The data, which are automatically generated when someone connects to a server with a view to downloading an article, includes more information than only the sole fact of a download, for instance the address of the computer from which the download was requested. This information could be used to establish crude profiles of the users of that document and these profiles could be even more precise in a closed system—for instance in Atkinson’s (1996, 256) “control zone”—where each user would be identified when logging in. In this online environment, the use of all objects could be tracked and even be differentiated by user group. We would be in a position to answer questions like, which items have been read (or at least retrieved) by which groups of readers (e.g., members of academic departments in the field), which have been read by scholars in other fields, by students, by the public at large? With a view to indicating something about the quality of a piece, one might envisage even differentiating between the standing of a scholar: if an article is widely read by those at the top of the hierarchy, this might be taken as an indicator for quality.

### ***Food for discussion***

All of these new tools for scientometrics are not without old (in the case of citation counts) or new (in the case of access counts and use-tracking) problems.

- (1) To begin with, it still remains dubious whether **counting citations or hits** does indeed say something important about the quality or performance of a scholar (e.g. Remler 2000). This is a long-standing debate (with regard to citations) and relates, at least in some academic fields, to the very heart of the academic culture. I think that dispute could be settled by an approach that does not put all the emphasis on either soft, qualitative criteria or on seemingly hard data generated by citation (or, now, hit) counts. As long as all assessors, committees and head of departments are aware of the strengths and weaknesses of both approaches and not one of both is reified, we shall be on the safe side. This is, however, not the case yet.
- (2) As regards access counts, in particular, there are still a number of technical problems to be solved before access figures would be of any use at all. It is relatively easy to **stimulate hits** on a web page (by asking friends and colleagues to visit it) or even generate them yourself, mechanically. However, if the practice of measuring web hits becomes important I am convinced that there will be ways to study adequately patterns of hits that represent real interest, and to extract useful information from these observed patterns (Ullman 1996).

(3) Furthermore, **not all access is alike**. A quick visit to an online article (just looking at the abstract and then deciding that it will not be used) would have to be differentiated from a long look at the full text or a download (Rohe 1998). This problem is similar to the known question of what should count as a real or countable “citation” (self-citations; reference to an example of “bad” science; straw man citations etc.). However, with certain limitations, it would be technically feasible to distinguish between various hits, taking into account time and action.

(4) While counting anonymous hits does not involve a **danger for privacy**, use tracking has such a potential. While the use tracking computers would have to be in a position to match use and identity to make use tracking possible, it would be necessary to prohibit these sensitive data from being disclosed. They should only be presented in anonymous and cumulative form. However, even if this were the case, it might still be possible to find out exactly who read what on an individual level since there may only be very few people of a certain category. While this problem is also present in today’s anonymous peer reviewing (the reviewer can often guess who the author is) generalised use tracking may be used to make much better guesses, in particular if data mining techniques are applied.

## Conclusions

The digital environment makes possible, at least in principle, that scientometrics replaces expert judgement on a case-by-case basis in evaluating the merits of a paper. It has been argued that citation ranking and expert judgement of scholars are highly correlated (So 1998). If this scenario becomes reality, the slogan “publish or perish” would be changed to “get hit or get out”. While counting papers is certainly no perfect indicator, the web allows us to measure readership (Ullman 1996). In the long run, citation ranking and use tracking have the potential to fully replace ex ante quality control. In the medium run, they might be a useful add-on to current mechanisms. However, we should be aware that there a number of technical and organisational problems to overcome first.

## References

- Alton-Scheidl, R., Schmutzer, R., Sint, P.-P. and Tscherteu, G. (Eds), 1997, *Voting, Rating, Annotation – Web4Groups and other Projects: Approaches and First Experiences*, Wien/München: R. Oldenbourg.
- Atkinson, R., 1996, Library Functions, Scholarly Communication, and the Foundation of the Digital Library: Laying Claim to the Control Zone, *The Library Quarterly* 66(3), 239-265.
- Atkinson, R., 2000, A rationale for the redesign of scholarly information exchange, *Library Resources and Technical Services* 44(2), 59-69.
- Cameron, R. D., 1997, A Universal Citation Database as a Catalyst for Reform in Scholarly Communication, *First Monday* 2(4) <[http://www.firstmonday.dk/issues/issue2\\_4/cameron/index.html](http://www.firstmonday.dk/issues/issue2_4/cameron/index.html)>.
- Davenport, E. and Cronin, B., 1990, Hypertext and the conduct of Science, *Journal of Documentation* 46(3), 175-192.
- Fröhlich, G., 1999, Das Messen des leicht Meßbaren. Output-Indikatoren, Impact-Maße: Artefakte der Szientometrie?, in: Becker, J. and Göhring, W. (Eds): *Kommunikation statt Markt – Zu einer alternativen Theorie der Informationsgesellschaft; GMD Report 61*, Sankt Augustin: Gesellschaft für Mathematik und Datenverarbeitung, 27-38 <<http://www.gmd.de/publications/report/0061/>>.
- Hibbitts, B. J., 1996, Last Writes? Reassessing the Law Review in the Age of Cyberspace, *New York University Law Review* 71(June), 615-688 <<http://www.law.pitt.edu/hibbitts/lastrev.htm>>; also published in: online original February 5, 1996.
- LaPorte, R. E., Marler, E., Akazawa, S., Gamboa, C., Shenton, C., Glosser, C., Villasenor, A. and Maclure, M., 1995, The death of biomedical journals, *British Medical Journal online* 310(May), 1387-1390 <<http://www.bmj.com/cgi/content/full/310/6991/1387>>.
- Nathenson, C., 2001, Electronic Publishing and the Future of Humanities Scholarship, *Trans, Internet-Zeitschrift für Kulturwissenschaften* (10) <<http://www.inst.at/trans/10Nr/nathenson10.htm>>.
- Nentwich, M., 2003, *Cyberscience: Research in the Age of the Internet*, Vienna: Austrian Academy of Sciences Press <<http://hw.oeaw.ac.at/3188-7>>.
- Nentwich, M., 2004, Quality control in academic publishing: challenges in the age of cyberscience, *Poiesis & Praxis. International Journal of Ethics of Science and Technology Assessment* 3, Online First <<http://springerlink.metapress.com/app/home/contribution.asp?wasp=c5v8d08ekg7jrv4b28f3&referrer=parent&backto=issue,4,14;journal,1,8;linkingpublicationresults,1:109376,1>>.
- Odlyzko, A. M., 1994, Tragic Loss or Good Riddance? The Impending Demise of Traditional Scholarly Journals, *Journal of Universal Computer Science* 0(0), 71-122 (quoted according to manuscript) <[http://www.jucs.org/jucs\\_0\\_0/tragic\\_loss\\_or\\_good](http://www.jucs.org/jucs_0_0/tragic_loss_or_good)>.
- Remler, A., 2000, Lässt sich wissenschaftliche Leistung messen?, *Telepolis*, 24.10. <<http://www.heise.de/tp/deutsch/special/auf/8976/1.html>>.
- Rohe, T. A., 1998, How Does Electronic Publishing Affect the Scholarly Communication Process?, *Journal of Electronic Publishing* 3(3) <<http://www.press.umich.edu/jep/03-03/rohe.html>>; also published in: Paper at the Faxon Institute Colloquium "Electronic Publishing and the Scholarly Communication Process, 7-8 January 1998.
- Rost, M., 1998, Elektronische Foren als Medien wissenschaftlicher Diskurse – Oder: Warum elektronische Foren in der vorliegenden Form nicht für den wissenschaftlichen Diskurs geeignet sind, *Telepolis Aktuell*, 1998-02-06 <<http://www.heise.de/tp/deutsch/inhalt/on/1389/1.html>>.
- So, C. Y. K., 1998, Citation ranking versus expert judgment in evaluating communication scholars: effects of research speciality size and individual prominence, *Scientometrics* 41(3), 525-333.
- Tomlins, C. L., 1998, The Wave of the Present: The Printed Scholarly Journal on the Edge of the Internet, *Journal of Scholarly Publishing* 29(April), 133-150.
- Ullman, J. D., 1996, Research Publication Modes Need to be Reengineered – A discussion, *Computing Research News*, July <<http://www-db.stanford.edu/~ullman/pub/nopaper.html>>.
- Varmus, H. and et al., 1999, *PubMed Central: An NIH-Operated Site for Electronic Distribution of Life Sciences Research Reports*; Last update: 1999-08-30 [Accessed on: 02-02 2000] <<http://www.nih.gov/about/director/pubmedcentral/pmcprint.htm>>.
- Wouters, P. F., 1999, *The Citation Culture*, Ph.D. thesis, Faculteit der Scheikunde, University of Amsterdam <[http://www.niwi.knaw.nl/en/nerdi2/group\\_members/paul/publ/citation/toonplaatje](http://www.niwi.knaw.nl/en/nerdi2/group_members/paul/publ/citation/toonplaatje)>.