OpenAccess Statistics: Alternative Impact Measures for Open Access documents? An examination how to generate interoperable usage information from distributed Open Access services

Ulrich Herb
Saarland University and State Library (Germany)

Abstract

Publishing and bibliometric indicators are of utmost relevance for scientists and research institutions as the impact or importance of a publication (or even of a scientist or an institution) is mostly regarded to be equivalent to a citation-based indicator, *e.g.* in form of the Journal Impact Factor or the Hirsch-Index. Both on an individual and an institutional level performance measurement depends strongly on these impact scores. This contribution shows that most common methods to assess the impact of scientific publications often discriminate Open Access publications – and by that reduce the attractiveness of Open Access for scientists. Assuming that the motivation to use Open Access publishing services (*e.g.* a journal or a repository) would increase if these services would convey some sort of reputation or impact to the scientists, alternative models of impact are discussed. Prevailing research results indicate that alternative metrics based on usage information of electronic documents are suitable to complement or to relativize citation-based indicators. Furthermore an insight into the project OpenAccess-Statistics OA-S is given. OA-S implemented an infrastructure to collect document-related usage information from distributed Open Access Repositories in an aggregator service in order to generate interoperable document access information according to three standards (COUNTER, LogEc and IFABC). The service also guarantees the deduplication of users and identical documents on different servers. In a second phase it is not only planned to implement added services like recommender features, but also to evaluate alternative impact metrics based on usage patterns of electronic documents.

Scientific Publishing, Scientometrics, Impact Metrics, Performance Measurement, Alternative Impact Metrics, Usage Information, Statistics

Published in:
L'information scientifique et technique dans l'univers numérique. Mesures et usages.
Edited by Chérifa Boukacem-Zeghmouri
ISBN: 978-2843651236
L'association des professionnels de l'information et de la documentation L'association des professionnels de l'information et de la documentation, ADBS, Paris, 2010

1 Scientific Publishing, Publication Impact, Scientific Capital, Reputation and Open Access

"Publish or perish": a scientist's career will fail if he does not publish. To be more precise: it will fail if he does not publish in the right contexts. But what makes a journal or some conference proceedings the right place to submit papers? Of course it is the impact that these journals, proceedings or publishers assign to a scientist. The higher the impact of such a publication is the better are the professional prospects of a scientist. The impact itself is mostly defined by citation indices, which may be assigned to a journal (as the Journal Impact Factor JIF) or to a scientist (as the Hirsch-Index or h-Index).

1.1 Citation-based Indicators: Journal Impact Factor and h-Index

1.1.1 Journal Impact Factor

The JIF of a journal Y is calculated as follows: *In year X, the impact factor of a journal Y is the average number of citations to articles that were published in Y during the two years preceding X.*

Although the inventor of the JIF, Eugene Garfield, asseverates: „We never predicted that people would turn this into an evaluation tool for giving out grants and funding" (MONASTERSKY, 2005, p. A 12) the JIF turned into a yardstick for the evaluation of scientific outcomes and the quality of publications.

The JIF is fiercely critized (CAMPBELL, 2008 ; DONG, LOH & MONDRY, 2005 ; SEGLEN, 1997 ; SEGLEN, 1998): The scope of the utilized publication data (the Web of Science WoS respectively the Journal Citation Reports JCR) is restricted and more or less arbitrarily defined. Only the Thomson Scientific's Institute for Scientific Information (ISI), the producer of the databases, decides which publication or journal is included and thereby indexed and which is not.

Additionally the JIF shows several disciplinary biases: The two years span discriminates publications from disciplines where information and knowledge have a lifespan that lasts longer than two year, *e.g*. Humanities, Social Sciences or Mathematics. Also the exclusion of a whole bunch of document types (monographs, proceedings, grey literature, etc.) disadvantages systematically disciplines that prefer other publication types than journals and advantages Science, Technology und Medicine. Furthermore journals in Non-English language are underrepresented in the JCR, so they can hardly reach high JIF-scores.

Surprisingly the JIF is not only used to rank journals, but also - against its logic - to rank scientist. These days it is quite common to calculate some sort of JIF for single scientists by multiplying their number of articles in journals X, Y, Z with the JIF-scores these journals, adding these results and dividing the sum by the overall number of articles a candidate published. Nevertheless the JIF does not give any testimony about the quality of a scientist or a scientist's publication. Within several contexts it is proven that in most cases a small number of highly cited articles produces a high JIF-score for the whole journal. For biochemical journals SEGLEN (1997 ; 1998) found out that 15% of the articles generated 50% of a journal's citations and that 50% of the articles generated 90% of a journals citations. CAMPBELL (2005) reports that 25% of the articles published in Nature in 2002 and 2003 generated 89% of the citations in these years, thereby they also made a great contribution to the JIF-score of 32.2 for the year 2005.

But not only the JIF-formula and the scope of the JIF are criticized, the crux lies in its interpretation which mostly occurs as a confoundation of popularity and quality.

1.1.2 h-Index

Unlike the JIF the *h-Index* is not related to journals but to authors: *A scientist has index h if h of N papers have at least h citations each, and the other (N − h) papers have less than h citations each.*

This means an author has an h-index of 8, if he published 8 articles that were cited at least 8 times. As it focuses on authors instead of journals the h-index seems to be more appropriate to assess scientists than the JIF. The h-index reveals a levelling but ambivalent nature: the citation count of one single highly-cited publication can not distort a scientist's score, but this peculiarity also depreciates innovative concepts.

Although its calculation is not by definition fixed to a database as the JCR a bigger part of the JIF-critique also applies to the h-index. Exemplary the neglect of many document types, the handling of documents in Non-English language and of documents from multiple authors are to be mentioned. Besides it would be very important to identify authors of publications accurately - unfortunately at that point most of the databases which are used for the calculation of the h-index are not faultless. Furthermore the h-index of a scientist depends on many context variables like his age or his discipline that make it hard to consider h-index scores comparable.

There a several suggestions to improve the h-index. For example the age bias shall be compensated by the so called normalized h-index: the quotient from h-index and a person's productive period (*e.g.* the years elapsed since the first publication). But one main question is also still not answered for the h-index: Does it measure popularity or quality?

At least a scientist's h-Index seems to correlate with the likelihood of a promotion. JENSEN, ROUQUIER & CROISSANT (2008) investigated for scientists from the Centre national de la recherche scientifique (CNRS) whether their scores according to several bibliometric indicators could predict their careers. Among the considered indicators the h-index was the best one to predict the career of a researcher – even though only for 48% of the persons a correlation between high h-index and career advancement could be found. Scientist could feel happy, if everyone would consider bibliometric indicators as differentiated as Jensen and his colleagues: "a 'mechanical Objectivity' procedure, which ranks candidates by their h would disagree with actual promotions for half of the promoted people, a very significant difference." (JENSEN, ROUQUIER & CROISSANT, 2008, p. 477)

1.1.3 Open Access and citation-based indicators

Commonly used citation-based indicators provide some arguments pro Open Access: Scientific documents that can be used free of charge are significantly more often downloaded and cited than Toll Access documents are (HARNAD & BRODY, 2004 ; LAWRENCE, 2001). Moreover the frequency of downloads seems to correlate with the citation counts of scientific documents (BRODY, HARNAD & CARR, 2006).

Nevertheless there is lack of tools and indicators to measure the impact of Open Access publications. Especially documents that are self-archived on Open Access Repositories (and not published in an Open Access Journal) are excluded from the relevant databases (WoS, JCR, Scopus, etc.) that are typically used to calculate JIF-scores or the h-index.

Open Access Journals on the other hand may have a JIF-score and indeed some of them even have an impressive Impact Factor, *e.g.* PLoS Medicine with a score of 12.185. However Open Access Journals often are discriminated by the JIF-formula and the scope of the JCR:

- Since many Open Access journals are quite new, they are lacking the citation history a journal needs to be indexed by the JCR and to reach an attractive JIF-score.
- Open Access Journals are published above average in developing countries (HAIDER, 2005). Due its unbalance towards the English language these journals usually attain minor JIF-scores – if they are indexed at all by the JCR.
- Accordingly PACKER and MENEGHINI (2007) found out that the JIF-scores of journals from the so-called developed countries are significantly higher than the JIF-score of journals form the so-called developing countries.

Generally it can be assumed that Open Access services (no matter if in the form of journals or repositories) would benefit from alternative impact indicators. Having the critiques on the JIF and the h-index in mind one might expect that scientific publishing as a whole would benefit from such indicators.

Furthermore there are also claims for transparent indicators like *Open Access to Citation Data* (SUBER, 2007) or *Open Metrics*: information of an extraordinary importance for organizational evaluation or individual careers like citation indexes data should be calculated in a transparent matter and should be verifiable. As Rossner, Van Epps and Hill (editors at Rockefeller University Press) tried to check the citation data and JIF-scores of three of their journals and several competing journals (ROSSNER, VAN EPPS & HILL, 2007; ROSSNER, VAN EPPS & HILL, 2008) they found repeatedly errors within the data provided by Thomson Scientific's Institute for Scientific Information which produces the WoS and the JCR. They reasonably concluded: "Just as scientists would not accept the findings in a scientific paper without seeing the primary data, so should they not rely on Thomson Scientific's impact factor, which is based on hidden data." (ROSSNER, VAN EPPS and HILL, 2007, p. 1092).

This bunch of arguments (the insufficiencies of conventional citation-based impact indicators, their fragmented coverage of Open Access documents and the claim for Open Metrics) raises the question if alternative indicators already exist or if they can be shaped. Ideally these alternatives indicators could make Open Access publishing more attractive by assigning impact and scientific capital (HERB, 2010) to Open Access publications - ascertained by a procedure that is accepted by colleagues, evaluators, bureaucracy and review board for professorship applicants.

2 Impact indicators: A categorization

As a first step citation-based indicators can be compared with usage-based indicators.

Citation-based indicators

- are author-centred: they measure citations which are actions of authors

- show a time delay: at least on generation of publications has to pass until a citation-based indicator may measure the impact of a publication
- try to measure impact on a journal level or an author level, but not on an item (article, monograph, dataset etc.) level.

Usage-based indicators in contrast

- are reader-centred: they measure document usage which is an action of readers
- can measure on-the-fly and consecutive
- can describe the impact of a single item
- allow an automatic measurement.

Apparently citation-based indicators and usage-based indicators can be considered to measure the impact of scientific items in a complementary way.

As a second step BOLLEN, VAN DE SOMPEL, SMITH & LUCE (2005) add the poles *frequency* and *structure* to the poles *author* and *reader*.

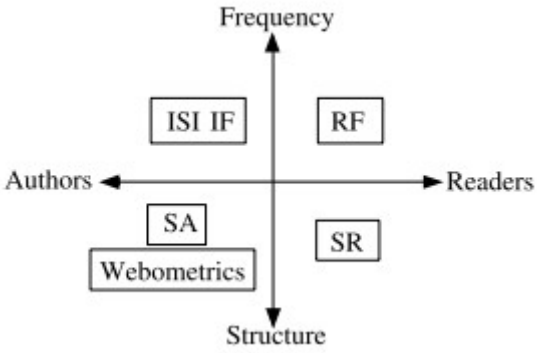

Figure 1, from Bollen, Van De Sompel, Smith & Luce, 2005, p. 1424

For instance the JIF (in Bollen's denomination ISI IF) lies in the quadrant of the author-centred frequency-metrics. Metrics that rely on networks of author actions (*e.g.* Google PageRank, citation graphs, hyperlink graphs or webometric indicators) are located in the structure/author-quadrant (SA) because they take not only heed of citation counts (as the JIF) but also citation or hyperlink patterns. In the upper right quadrant metrics that are based on the frequencies of document usage can be found, Bollen and his colleagues use the notion *reading factor* (RF) to embrace these concepts. The RF-quadrant is home of metrics that count the absolute frequency of document downloads or usage, as for example

- COUNTER (Counting Online Usage of Networked Electronic Resources)[1]
- LogEc[2] (the statistics module of the network RePEc[3])
- International Federation of Audit Bureaux of Circulations [4] (IFABC)

The *Code of Practice for Journals and Databases* (COUNTER, 2008) measures usage of documents on the aggregated level of journals. This information is predominately used by libraries to control the cost-effectiveness of their journal subscriptions. Since Open Access Journals can be used free of charge COUNTER is not prepared to measure their impact: For instance the *List of Internet Robots*[5] that COUNTER uses to eliminate non-human accesses on the articles may be useful if the documents are located on a publisher's server where they can not be indexed by search engines, but it is not appropriate to count hits on articles in Open Access Journals that are not only accessible for every person but also for every robot. Furthermore it generates only reports on the level of journals and therefore it is not very suitable to measure the impact of single objects. Unfortunately documents that are deposited on repositories are not at all considered by COUNTER. LogEC in fact measures document usage on the level of articles with the intent to measure the impact of single objects. Non-human hits are eliminated according to robots lists and with elaborated statistical procedures:[6]

For instance a host is considered a robot if the C-class net it belongs to accesses more than 10 % of the items in RePEc. Also the method of the IFABC measures the usage of single documents. It fits the requirements of the online advertisement industry which wants to monitor the usage of single websites in order to calculate the number of views embedded advertisement attained. DRIVER (2008, p. 131-135) gives an overview of the

[1] http://www.projectcounter.org
[2] http://logec.repec.org/
[3] http://repec.org/
[4] http://www.ifabc.org/
[5] http://www.projectcounter.org/r3/r3_K.doc
[6] http://logec.repec.org/about.htm

different procedures mentioned and their parameters. Some of the manifold and partly very problematic methodological and technical conditions and simplifications should be mentioned briefly. Although there are several proposals to measure the impact of online documents, we are far from *one accepted standard*. The models reveal considerable differences regarding the detection and elimination of non-human accesses and the definition of doubleclick-intervals. Moreover they neglect the context of the document usage. More importantly these models ignore also the detection of duplicate users and documents: To gather information on the context of document usage it would be necessary to detect what different documents user X downloaded from different servers. Besides it should be possible to sum up hits on different documents of the same content on different servers. On a very fundamental level it might be discussed what extent of similarity makes two different files the same document and how to distinguish different versions of a document (NISO, 2008).

Within the quadrant *Structure Reader* (SR) there are measures that consider the context of document usage like download graphs. Bollen and his colleagues carry out research in this area (BOLLEN, VAN DE SOMPEL & RODRIGUEZ, 2008 ; BOLLEN, VAN DE SOMPEL, HAGBERG & CHUTE, 2009 ; BOLLEN, VAN DE SOMPEL, SMITH & LUCE, 2005): They collected information about citations and usage of documents both in form of pure frequencies (of citation and usage) and in form of structural context information (as networks of document citations and document usage). This information was analysed with sociometric methods and techniques known from the network analysis. The rankings they produced were evaluated by scientists from different scientific communities with a suprising result: Some of the rankings based on the context of document usage echoed the scientists' preferences better than the JIF.

Admittedly Bollen's research happens under controlled laboratory conditions: As the researchers often use data from linkresolvers and aggregators it is not always certain whether an user only hit the abstract page or whether he really downloaded a paper. Additionally other things like handling double-click intervals, elimination of duplicate users or documents and the detection and elimination of non-human document accesses are mainly not a problem in Bollen's research environment.

3 Usage-based indicators: An assessment

Apparently alternative, usage-based impact measures can be designed. Despite the missing standardization usage-based indicators seem not only to predict the results of citation-based indicators (BRODY, HARNAD & CARR, 2006), they also express a distinct sort of impact (BOLLEN, VAN DE SOMPEL, HAGBERG & CHUTE, 2009) – but certainly these impact indicators have to be solidly evaluated. Analogue to the claim for *Open Access to Citation Data* also *Open Access to Usage Data* should be allowed, even to the point of licensing the usage data under a Creative Commons License.

The most promising procedure is designed by Bollen and his colleagues and it is also the most complex procedure. To test it in a true Open Access environment, including empirical noise that is unknown to Bollen's test bed, it needs a sophisticated infrastructure to generate and exchange interoperable usage information within a network of several different servers. For example this includes the logging of usage events on Open Access Repositories that are indexed by legions of robots and that contain multi-file documents and duplicate documents (maybe in different file formats). An infrastructure like that faces all the problems known from weblog analysis in digital libraries as reported for instance by JAMALI, NICHOLAS & HUNTINGTON (2005).

4 OpenAccess Statistics: An examination how to generate interoperable usage information from distributed Open Access services

In the project „Open Access Statistics (OA-S)"[7] (funded by the German Research Foundation DFG[8]) the project partners[9] built an infrastructure like that. OA-S was initiated by the Electronic Publishing working group of DINI[10] (Deutsche Initiative für Netzwerkinformation / German Initiative for Network Information) and has two tightly associated projects. While Open Access Statistics addresses usage description,

- Open Access Citation (or Distributed Open Access Reference Citation Services DOARC)[11] address the issue of tracking citations between electronic publications
- Open Access Network[12] intends to build a network of repositories and it will also bundle the results of Open Access Citation and OA-S in one user interface[13]. It also offers services for Open Access Citation and

[7] http://www.dini.de/projekte/oa-statistik/english
[8] http://www.dfg.de
[9] Georg-August Universitaet Goettingen (State- and University Library), Humboldt-University Berlin (Computer- and Mediaservice), Saarland University (Saarland University and State Library), and the University Stuttgart (University Library)
[10] http://www.dini.de
[11] http://doarc.projects.isn-oldenburg.de/

OA-S, *e.g*. the deduplication of documents which is based on an asymmetric similarity of full text documents.

OA-S implemented a network to collect and exchange usage information between different services and to process this information according to the standards of COUNTER, LogEc and IFABC. Additionally OA-S outlined additional services for repositories based on usage information and it developed implementation guidelines which make it easy for other services to join the OA-S network.

On the technical level the data providers (Open Access Repositories, licence controlling servers, linkresolvers) at the four partner institutions

- generate logs about document usage
- pseudonymize user information (*e.g*. IP-addresses)
- process usage information (add unique document ID, transforms data into OpenURL ContextObjects, …)
- transmit the information *via* OAI-PMH to the aggregation server (central service provider)

After collecting the usage events from each single data pool the central service provider process this data:

- it deduplicates documents: *e.g*. it sums up the hits on files with the same content on different servers
- it deduplicates users, so it is possible to create download graphs or to conduct clickstream analysis
- it processes the data according to the standards mentioned (including the removal of non-human accesses and considering standard-specific parameters like doubleclick spans)

After the calculation the usage data will be retransferred to the distributed services (the data providers) and to the Open Access Network service.

Data providers have to fulfil rather light-weight requirements to take part in the OA-S network: their web servers have to use a defined but easy to handle configuration (HERB et *al*., 2009), they must pseudonomyze user information and isolate the local document identifier and as a last step they have to offer the information as OpenURL-ContextObjects containers (with the elements referent, referring entity, requester, servicetype, resolver and referrer) *via* an OAI-PMH-interface to the service provider or aggregator service. DSpace- or OPUS-repositories may even use modules developed by OA-S, other products can easily be configured to be OA-S-ready.

The usage data produced by OA-S may be used

- from an *user perspective* as a criterion to estimate the relevance of a document (*e.g*. rankings)
- from an *author perspective* as an indicator for the dissemination of a concept
- from a *service provider* perspective:
    * as additional metadata for search engines, databases …
    * as a recommender service
- from a *repository perspective*
    * as a recommender service
    * as additional metadata for users

Some lessons OA-S learned by now are that:

- Linkresolver logs are hard to integrate in the framework. Some services (OVID) do not offer suitable information while the information form other services (SFX) seem very heterogeneous.
- The deduplication of documents appears very difficult for several reasons. For instance:
    * A document may have more than one ID or even more than one persistent identifier due to multiple deposits on different repositories
    * Two documents with exactly the same content may use different sorts of persistent identifiers. The formal publication in a journal may have a persistent identifier in form of a DOI, while the postprint in a repository has a persistent identifier in form of an URN.
    * A given document may have several splash pages on different servers pointing to one single file on one server due to metadata harvesting.

By now OA-S strives for a second funding phase. Some of the core points for OA-S II will be

- the internationalisation of the project,
- the standardization of indicators that are based on the absolute frequency of document usage,
- the integration of new contributing services/ data providers (in form of journals or repositories)
- the evaluation of indicators that are more complex (mostly using techniques of usage data network analysis) than pure usage frequencies of documents and
- the implementation of added-value services for repositories based on usage data.

Especially internationalisation and standardization need an intense exchange of information with other projects tackling related issues as SURFsure, COUNTER, PIRUS, NEEO, PEER or OAPEN and Knowledge Exchange,

---

[12] http://www.dini.de/projekte/oa-netzwerk/

[13] http://oansuche.open-access.net/findnbrowse/pages/start.faces

the cooperation of Denmark's Electronic Research Library (DEFF), the German Research Foundation (DFG), the Joint Information Systems Committee (JISC) and the SURFfoundation.

Literature

BOLLEN, J., VAN DE SOMPEL, H., (2008). "Usage Impact Factor: The Effects of Sample Characteristics on Usage-Based Impact Metrics". *Journal of the American Society for Information Science and Technology*, vol. 59, n° 1, p. 136-149.
DOI: 10.1002/asi.20746

BOLLEN, J., VAN DE SOMPEL, H., SMITH, J. A., LUCE, R., (2005). "Toward alternative metrics of journal impact: A comparison of download and citation data". *Information Processing and Management*, vol. 41, n° 6, p. 1419-1440.
DOI: 10.1016/j.ipm.2005.03.024

BOLLEN, J., VAN DE SOMPEL, H., HAGBERG, A., CHUTE, R., (2009). "A principal component analysis of 39 scientific impact measures". *PloS one*, vol 4, n° 6, e6022.
DOI: 10.1371/journal.pone.0006022

BRODY, T., HARNAD, S., CARR, L., (2005). "Earlier Web Usage Statistics as Predictors of Later Citation Impact". *Journal of the American Association for Information Science and Technology*, vol. 57, n° 8, p. 1060-1072.

CAMPBELL, P., (2005). "Not-so-deep impact". *Nature*, vol. 435, n° 77045, p. 1003-1004.
DOI: 10.1038/4351003a

CAMPBELL, P., (2008). "Escape from the impact factor". *Ethics in Science and Environmental Politics,* vol. 8, n° 1, p. 5-7.
DOI: 10.3354/esep00078.

COUNTER, (2008). *The COUNTER Code of Practice - Journals and Databases*: Release 3 Edinburgh, United Kingdom.
www.projectcounter.org/r3/Release3D9.pdf

DONG, P., LOH, M., MONDRY, A., (2005). "The "impact factor" revisited". *Biomedical Digital Libraries*, vol. 2, n° 7.
DOI: 10.1186/1742-5581-2-7

DRIVER, (2008). *DRIVER Guidelines 2.0 - Guidelines for content providers exposing textual resources* via *OAI-PMH.*
www.driver-support.eu/documents/DRIVER_Guidelines_v2_Final_2008-11-13.pdf

Haider, J., (2005). The Geographic Distribution of Open Access Journals.
http://www.soi.city.ac.uk/~ch696/icml9_abstract_haider.pdf

HARNAD, S., BRODY, T., (2004). "Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals". *D-Lib Magazine*, vol. 10, n° 6.
DOI: 10.1045/june2004-harnad

HERB, U. (2010). "Sociological implications of scientific publishing: Open access, science, society, democracy, and the digital divide". *First Monday*, vol. 15, n° 2.
http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2599/2404

HERB, U., Mittelsdorf, B., Metje, D., Schulze, M., Horstmann, W., Dobratz, S., Scholze, F., (2009). "Nutzungsstatistiken elektronischer Publikationen. Deutsche Initiative für Netzwerkinformation DINI".
http://edoc.hu-berlin.de/series/dini-schriften/2009-11/PDF/11.pdf

JENSEN, P., ROUQUIER, J.-B., CROISSANT, Y., (2009). "Testing bibliometric indicators by their prediction of scientists promotions". *Scientometrics*, vol. 78, n° 3, p. 467-479.
DOI: 10.1007/s11192-007-2014-3

LAWRECE, S., (2001). "Free online availability substantially increases a paper's impact". *Nature*, vol. 411, n° 6837, p. 521.
DOI: 10.1038/35079151

MONASTERSKY, R., (2005). "The Number That's Devouring Science". *The Chronicle of Higher Education*, vol. 52, n° 8, p. A 12.
http://chronicle.com/free/v52/i08/08a01201.htm

NISO (2008): *Journal Article Versions (JAV)*: *Recommendations of the NISO/ALPSP JAV Technical Working Group*.
www.niso.org/publications/rp/RP-8-2008.pdf

PACKER, A. L., MENEGHINI, R., (2007). “Learning to communicate science in developing countries”. *Interciencia*, vol. 32, n° 9, p. 643-647.
www.interciencia.org/v32_09/643.pdf

ROSSNER, M., VAN EPPS, H., HILL, E., (2007). “Show me the data”. *The Journal of Cell Biology*, vol. 179, n° 6, p. 1091-1092.
DOI: 10.1083/jcb.200711140

ROSSNER, M., VAN EPPS, H., HILL, E., (2008). “Irreproducible results: a response to Thomson Scientific”. *The Journal of Experimental Medicine*, vol. 205, n° 2, p. 260-261.
DOI: 10.1084/jem.20080053

SEGLEN, P. O., (1997). “Why the impact factor of journals should not be used for evaluating research”. *British medical journal*, vol. 314, n° 7079, p. 498–502.
www.bmj.com/cgi/content/full/314/7079/497

SEGLEN, P. O., (1998). “Citation rates and journal impact factors are not suitable for evaluation of research”. *Acta Orthopaedica*, vol. 69, n° 3, p. 224 - 229.
DOI: 10.3109/17453679809000920

SUBER, P., (2007). “Why we need OA to citation data”. *Open Access News. News from the open access movement.*
www.earlham.edu/~peters/fos/2007/12/why-we-need-oa-to-citation-data.html