Measuring Research Impact in an Open Access Environment

by FRANK SCHOLZE

ABSTRACT

This paper focuses on electronic publication impact as a limited, but rather well defined sub-field of research impact. With Open Access, a much bigger corpus of data has become available for statistical analysis. Publication impact can be measured by author- or reader-generated data. Author-generated data would be citations. Reader-generated data would be usage. Usage data can be collected through webserver or linkresolver logs. It has to be normalized in order to be shared and analysed meaningfully. The paper presents current initiatives and projects aiming to provide a suitable infrastructure, including publisher data (COUNTER/SUSHI) and data collected from Open Access repositories (using OAI-PMH and OpenURL ContextObjects). Citation and usage data can be analyzed quantitatively or structurally. These new metrics can enhance or complement existing metrics like the Journal Impact Factor (JIF). Services like decision support systems for collection management or recommender systems can also be built on this metrics.

WHY MEASURE RESEARCH IMPACT?

Although this is a highly political question it has been a long tradition to compare, measure (and honour) scientific achievement. Assessment and evaluation of research are important for a number of reasons, among them appointment decisions, funding decisions, the need to monitor trends and the need to prioritize activities and attention. As sensible as these reasons might be there is a considerable amount of uneasiness among scientists about being assessed and evaluated. The danger of comparing apples and pears seems just too high for some of them. As evaluation is to a large extent driven from outside the scientific community, just not to play the game seems not enough. As Alan Gilbert, President of the University of Manchester put it in a recent Nature Article "Rankings are here to stay, and it is therefore worth the time and effort to get them right."([Butler, 2007](#)).

However impact and research are very broad and fuzzy concepts. It is obvious that it needs several levels of abstraction in order to describe, assess and measure research impact in a meaningful way. There is a range of qualitative and quantitative methods deployed in social sciences to determine impact ([Creswell, 2003](#)). Although there is still an argument among social scientists about the pre-eminence of qualitative or quantitative methods, it is justified to say that in measuring research impact approaches from both domains are complementary. Qualitative approaches would include any voting or reviewing systems. Publications are the quantifiable output of the research process. It is therefore one manifest option to build quantitative metrics onto publications in order to measure research impact. There are other choices for collecting quantitative data on research (like third party funding, cooperation projects, licenses, start-ups, doctoral students etc.), but this paper will focus on publications. So it reduces the complexity from research impact to publication impact. It further narrows its scope by focusing on electronic publications as these can be handled with automated procedures, which scales up better for large amounts of data. Impact implies an interaction, so mere numbers of publications would not be sufficient. It is rather citations and usage that have to be regarded as indicators of effect.

OPEN ACCESS AND METRICS

Open access (OA) is the immediate, free and unrestricted online access to digital scholarly material, primarily peer-reviewed research articles in journals. There are two main currents in the open access movement:

1. In OA self-archiving (sometimes known as the "green" road to Open Access), authors publish in a traditional subscription journal, but in addition make their articles freely accessible online, usually by depositing them in an institutional or central repository.

2. In OA publishing (sometimes known as the "gold" road), authors publish in open access journals that make their articles freely accessible online immediately upon publication.

Both approaches are complementary and in relation to impact metrics have two effects. On the one hand, more and more publications become freely available. Therefore the possibility to collect and process quantitative data on electronic publications becomes more widespread as citations and usage of Open Access publications can become freely available as well. Policies of Open Access repositories should at least express an opinion on the availability of

citation and usage data and it has already been proposed to enhance the OpenDOAR policies tool accordingly in order to raise awareness among repository managers.

On the other hand Open Access acts as a catalyst for thinking about the possibility of constructing new indicators to measure different aspects of research impact and to enhance and complement existing metrics (Scholze & Dobratz, 2006). The fast-changing nature of scholarly communications makes us aware that traditional metrics might not be exclusively sufficient to describe research impact.

JOURNAL IMPACT FACTOR

The Journal Impact Factor (JIF), often abbreviated to IF, is a measure of citations to scientific journal literature. It is frequently used as a proxy for the importance of a journal to its field and may be the most well-known and prominent impact measure applied to publications. The impact factor for a journal is calculated based on a three-year period, and can be considered to be the average number of times published papers are cited up to two years after publication. For example, the 2006 impact factor for a journal would be calculated as follows:

The number of citations articles published in a Journal X in 2004-5 received from all articles in all Journals indexed in 2006 d*ivided by* the number of articles published in a Journal X in 2004-5

So the Journal Impact Factor is just a mean 2-year citation rate which gives a fair approximation of journal 'status' but should not be used to rank authors, departments, institutions, regions, nations, etc. as it does not give any evidence about their performance. However it is used in many cases now in tenure, promotion and other evaluation procedures. The very term Journal Impact Factor is a good indicator that it is simply not suited for such uses and it is its misuse as a general Impact Factor that has caused much criticism and misunderstanding (Seglen, 1997; Dong, Loh & Mondry, 2005). Even Eugene Garfield the creator of the JIF argues heavily against such misuse: "We never predicted that people would turn this into an evaluation tool for giving out grants and funding."(Monastersky , 2005).

This ambivalent view of the Journal Impact Factor is duly reflected in a recent survey commissioned by the UK Serials Group (Fig. 1). About 1400 academics were asked about their opinions on the JIF as a measure for the performance of a Journal and as a measure for the assessment of individual scholars. They largely agreed that the JIF is a valid measure of the quality of a Journal (47% either agreed or strongly agreed) but maintained that it is too much used for assessing individual scholars (62% either agreed or strongly agreed).
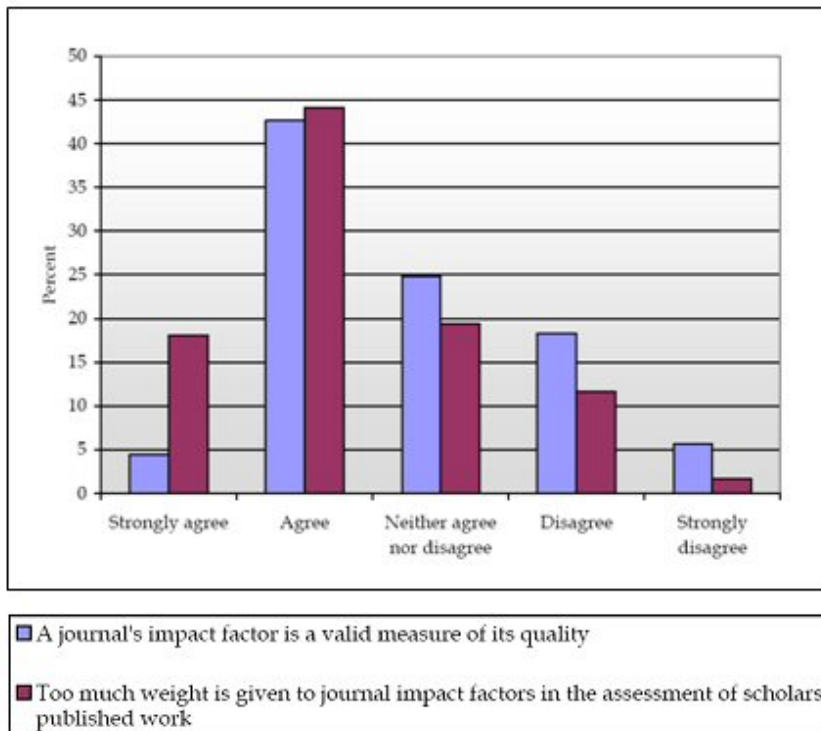

Fig. 1: Academic authors' views on the value and use of Journal Impact Factors (Shepherd, 2007)

OTHER APPROACHES

What else is there except the JIF? Social network analysis developed systematically in the 1950s and 60s has emerged as a key technique in modern sociology, anthropology, sociolinguistics, geography, social psychology, communication studies, information science, organizational studies, economics, and biology (Freeman, 2004). It has developed a range of metrics (among them betweenness, centrality, clustering or eigenvector centrality) which can be applied to publication networks of usage and citations as well. The most well known algorithm out of this range, due to its widespread use, is page rank.[1] It can be used to calculate the relative weight of reciprocal voting of nodes in a network. It can be applied to any collection of entities with reciprocal references. The formula uses a model of a random surfer who gets bored after several clicks and switches to a random page. The page rank value of a page reflects the chance that the random surfer will land on that page by clicking on a link. It is a elegant variant of the eigenvector centrality measure, as its values are fast to approximate and it gives good results (Google being the proof for this statement). Related approaches in network analysis include authority and hub values - HITS algorithm (Kleinberg, 1998). Although many of these algorithms were designed to analyse the WWW (therefore the term webometrics was coined) they can be applied to publications networks as well. Their promise lies in their ability to express relative values out of structural characteristics from within a network. The postulate is that within the structure of a network like citation and usage of publications is a wealth of information about the relative influence of individual articles, journals and scholars, and also about the patterns of relations among academic disciplines. Bollen and Van de Sompel were among the first to describe the application of social network analysis and webometrics to publication networks in order to enhance and complement existing metrics in this field (Bollen, Van de Sompel & Luce, 2005; Bollen, Rodriguez & Van de Sompel, 2006). They also introduced a taxonomy to put existing and new approaches to metrics into context, taking author- (i.e. citations) vs reader- (i.e. usage) generated data and structural vs frequentist methods as discriminating elements (Fig. 2).
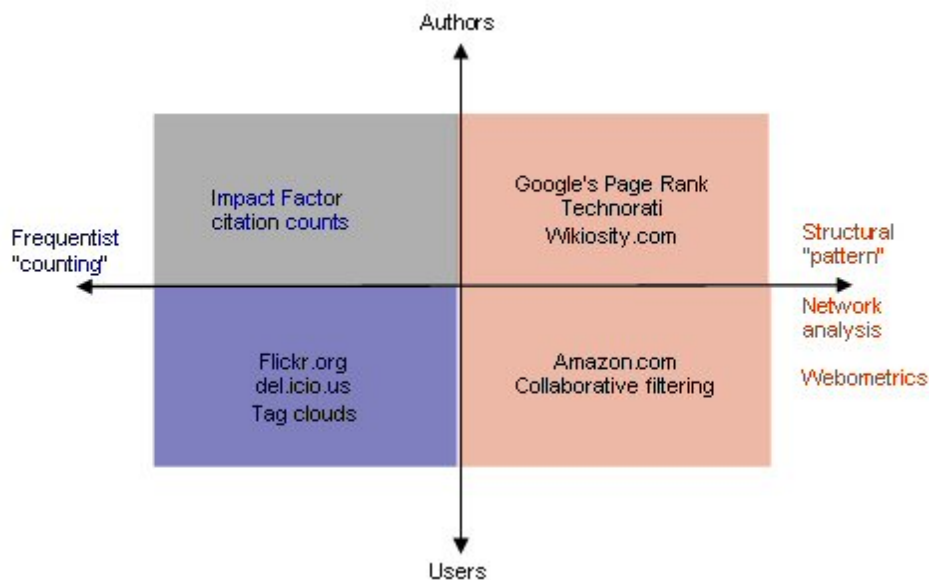

Fig. 2: A Taxonomy of Metrics (Bollen & Van de Sompel, 2005)

The examples given show that network analysis algorithms not only function as a tool for analysing networks built in the past, but also can act as part of recommender systems to make decisions for the future based on this past information. Wikiosity, for example, is a recommender system for Wikipedia articles based on the structure of links in Wikipedia. Amazon uses a recommender system based on items bought or accessed by previous customers. The same can be done for publication networks. Recommendations for related publications can be based on usage data either from an online catalogue or a digital library or a broader publication network.

ELEMENTS IN MEASURING PUBLICATION IMPACT

This section briefly describes the different elements or layers needed to measure publication impact (Fig. 3).
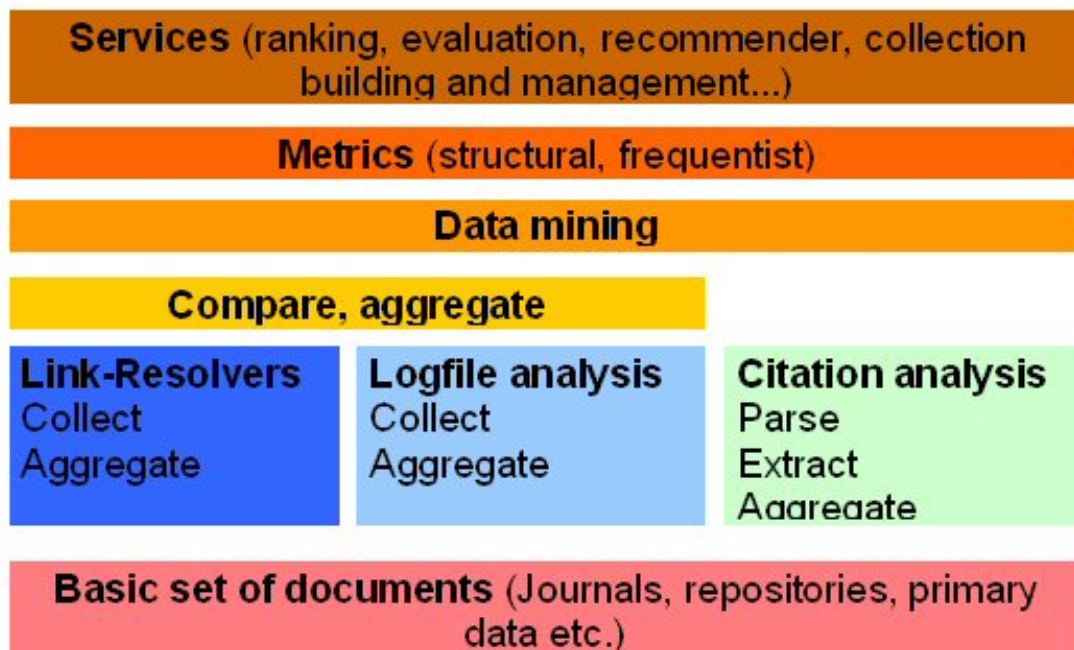

Fig. 3: Elements in measuring publication impact

As often in social sciences, it is crucial which basic set of data (in this case publications or documents) are taken into account and how representative they are. Numerous choices are possible, the dataset(s) for the JIF being just one.[2] Open Access has made even a greater range of choices possible with more than 1000 repositories listed by OpenDOAR and more than 2900 journals listed by DOAJ. The next level is to look at author- or reader-generated data, i.e. citations or usage. For citations the most comprehensive source of data is still the Science Citation Index (SCI) by Thomson/ISI which has manually-acquired citation data back to 1900. Thomson begins to cover citations from and to OA content with the Web Citation Index (WCI). Related automated approaches based on autonomous citation indexing include Google Scholar, Citebase and Citeseer. SCOPUS from Elsevier, the comparatively new competitor to SCI/WCI, also includes OA content (Lawrence, Giles & Bollacker, 1999).

COLLECTING USAGE DATA

For collecting usage data about publications, two basic approaches are possible. One is based on logfile analysis and one based on link resolver logs. Client-side approaches like web bugs or pixel tags common in web page statistics are not sufficient for distributed publication networks where documents may exist in different versions, formats and multiple single files. Logs from repositories or journal sites, as well as from link resolvers, can be made accessible by standard technologies using the OAI Protocol for Metadata Harvesting (OAI-PMH) and OpenURL ContextObjects (Fig. 4). The basic architecture has been proposed by Bollen and Van de Sompel and can be expanded to include data from publisher sites available in a different XML-form called SUSHI (Bollen & Van de Sompel, 2006). At the moment, however, statistical data from publisher sites conformant to the COUNTER initiative is only available at the journal title level and not at the article level. Therefore aggregating data from these sources results in a much coarser granularity.
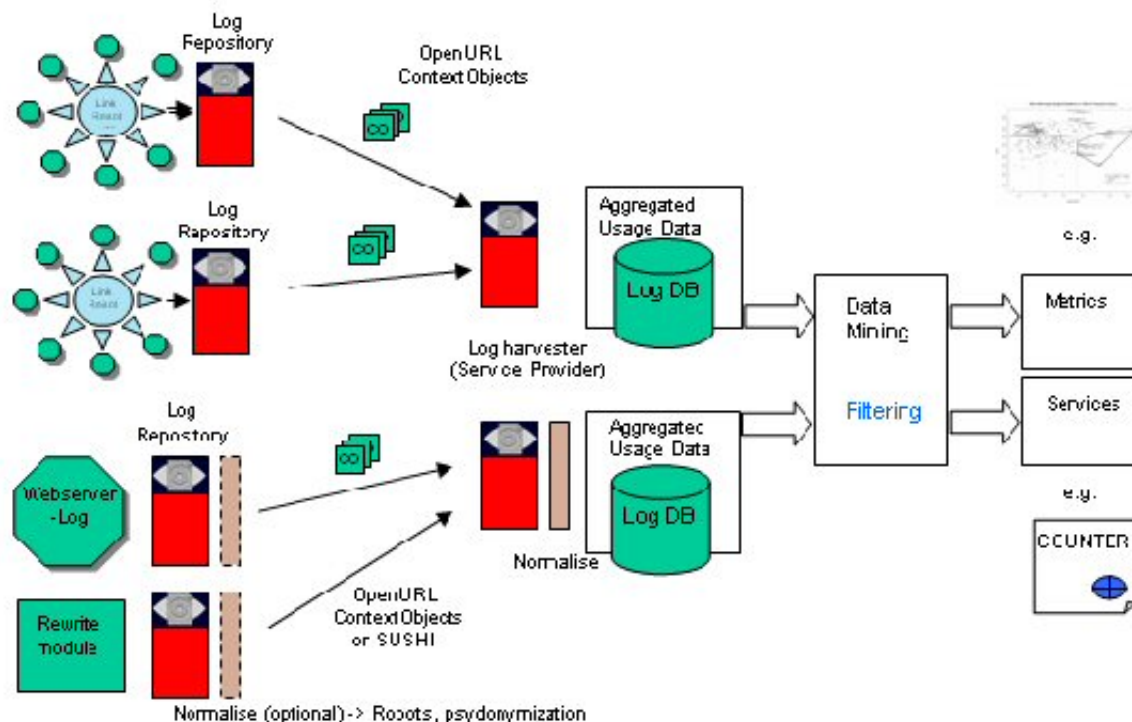
Fig. 4: An infrastructure for collecting usage data (Bollen & Van de Sompel, 2005)

Data aggregated from different sources has to be normalized, automated accesses from robots have to be tagged and duplicates have to be removed. This last point refers to publications and can be done based on persistent identifiers (like DOIs or URNs). In a broader context, the removal of duplicates will also have to rely on metadata-based heuristics. This means that duplicates are detected based on the comparison of ISSN-numbers, article titles or publication year or parts or combinations of these fields.

After these preliminary steps of data collection and preparation, data mining techniques can be invoked, that is the extraction of useful information from the large data sets collected in the preliminary phases. Metrics and services can now be based on this consolidated data. It is possible to calculate the whole series of metrics among them a usage page rank as well as a citation page rank. Usage page rank has been calculated with data from the Los Alamos National Laboratory (LANL), California State University and University of Texas in the bX project carried out by LANL and ExLibris (Bollen, Beit-Arie & Van de Sompel, 2005) and the MESUR project (currently carried out by LANL).

In the bX project, usage data at California State University was collected through an infrastructure of SFX link resolvers. Among a series of other investigations, page rank based on usage data was calculated and then correlated with the respective JIF on the journal title level for different disciplines (cf. Fig. 5 for Computer Science). Interesting clusters of titles showed a strong disparity between usage page rank and JIF. One group showed a strong dominance of the JIF with a comparatively low local usage (prestigious but not popular titles like the *Journal of Molecular Graphics and Modelling* in the upper left). The other group showed high local usage and a comparatively low JIF (popular but not prestigious titles like *Dr. Dobbs Journal* or *Communications of the ACM* in the lower right). This latter feature is especially important when, for example, it comes to cancellations due to budget cuts for example. It also shows that JIF as a single metric to distinguish a journal's impact is not sufficient as titles like *Communications of the ACM* are read more frequently and therefore inform and influence scholars and students much more than could be expected from their JIF. It is important to note that usage data was collected on the journal article level and aggregated to the title level at a later stage for reasons of comparison. With the granularity of data initially collected, it is possible to relate usage factors to individual scientists, institutes, faculties etc. depending on different needs and purposes for evaluation.
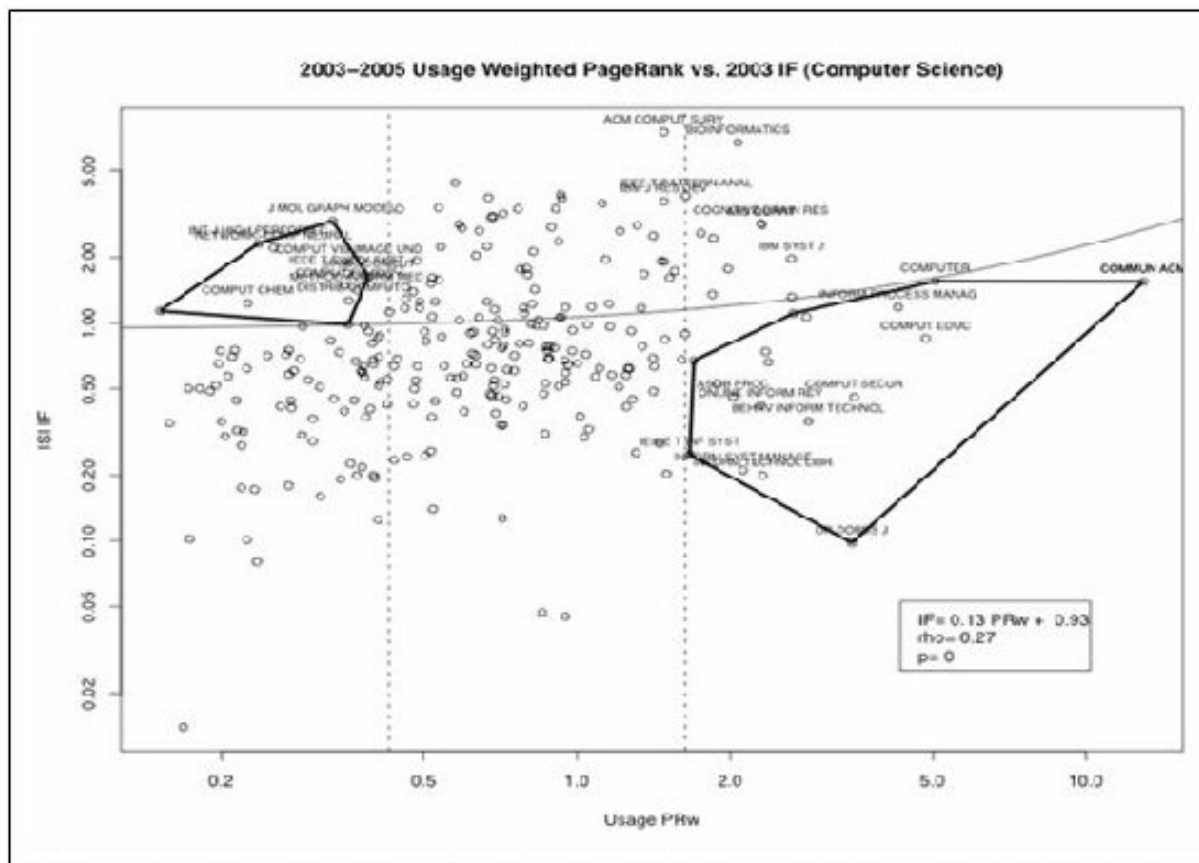
Fig. 5: Comparison of Journal Usage Page rank and Journal Impact Factor in Computer Science, California State University (Bollen, Beit-Arie & Van de Sompel, 2005).

The results from bX show that, as a by-product, collection management or decision support systems can be based on usage or citation data, particularly if they are correlated with pricing information. The service eigenfactor.org based at the University of Washington in collaboration with journalprices.com provides ranked information about price and value for thousands of scholarly periodicals. Eigenfactor uses an adapted form of page rank applied to citation information in order to calculate the importance of each journal.

CURRENT INITIATIVES

DINI, the Deutsche Initiative für Netzwerkinformation (German Initiative for Network Information), is a voluntary organisation recruiting from Computer and Media Centers as well as Libraries and roughly comparable to the Coalition for Networked Information (CNI) in the US. DINI has put forward a cluster of proposals to the DFG, the Deutsche Forschungsgemeinschaft (German Research Foundation) in order to build a network of certified open access repositories which is seen as the national input to the EU repository infrastructure project DRIVER.. In addition to this networking of repositories, one of the proposals aims to build an infrastructure to collect and aggregate usage data on a national level. Together with the LIBER Access Division and DRIVER, DINI intends to spread this approach across Europe. Currently a demonstrator of this infrastructure, and some of the related services, is being presented to the DFG. The partners in this project are Göttingen State and University Library, Stuttgart University Library, Computer and Media Service Humboldt University Berlin and the Saarbrücken State and University Library. The project will also take into account recent activities by the German collecting society for copyright charges (VG Wort) which has implemented a statistics-based payment to authors for electronic publications (METIS). This is a redistribution of charges collected from Libraries, CD and DVD manufacturers etc. as a compensation for making copies for private use. A co-operation is planned where the VG Wort is one of the service providers receiving aggregated usage data using OAI-PMH and OpenURL ContextObjects. The VG Wort applies different rules and normalizations to usage data based on the IFABC recommendations which differ for example from the COUNTER Codes of Practice. The activities in Germany and the UK are therefore focused on the infrastructure for collecting and aggregating high quality usage data. This has been underpinned by a workshop organized by the Knowledge Exchange initiative. In 6 challenging reports written as a result of the workshop, experts from 7 European countries make recommendations on the topics related to institutional repositories - among them a usage statistics report (KE, 2007). Knowledge Exchange is a co-

operative effort that intends to support the use and development of ICT infrastructure for Higher Education and research in Germany, the UK, Denmark and the Netherlands.

Complementary to the infrastructure oriented activities in Europe project MESUR will assess and validate a wide range of indicators and metrics for publication impact in order to gain insight into the relevance, validity and appropriateness of metrics for different purposes. At the moment, it is still not clear which metric or which combination of metrics is suitable for certain objectives. Usage-based metrics can express different facets of publication impact. Harnad proposed a related approach of testing and validating metrics with qualitative data from reviewers in the forthcoming parallel panel-based and metric UK Research Assessment Exercise in 2008 (Harnad, 2007).

As both the infrastructure for collecting and aggregating data, and the metrics for meaningful and expedient interpretation, grow more mature they will be powerful tools for the much-needed evaluation and assessment of research in the light of rapidly changing scholarly communications.

CONCLUSION

Scholarly evaluation will evolve as scholarly communication does. In the field of electronic scientific publications, usage data comes more and more into focus as a basis for metrics. Open Access acts as a catalyst for getting free access to more and richer data about publications as well. The infrastructure for collecting and aggregating usage data (developed largely at LANL) is conceptually available but has to be deployed and implemented in practice on a large scale. DINI, together with the LIBER Access Division and DRIVER, are trying to accomplish this for Europe.

Building on the tradition of network analysis, and advanced by its application in the WWW, a new range of structural metrics (among which page rank is the most prominent) is being tested for scholarly evaluation. These metrics can enhance and complement the omnipresent JIF. MESUR is investigating and validating metrics for different needs and purposes. New metrics and richer data will give us the opportunity "to get rankings right", tailoring them to specific objectives.

REFERENCES

Bollen, J. and H. Van de Sompel: "A framework for assessing the impact of units of scholarly communication based on OAI-PMH harvesting of usage information". Presentation delivered at CERN workshop on Innovations in Scholarly Communication (OAI4), Geneva, Switzerland, 2005. http://eprints.rclis.org/archive/00006076/

Bollen, J. and H. Van de Sompel: "An architecture for the aggregation and analysis of scholarly usage data". Proceedings of the 6th ACM/IEEE-CS joint conference on Digital libraries, 2006, pp. 298 – 307 http://doi.acm.org/10.1145/1141753.1141821

Bollen, J., Oren Beit-Arie and H. Van de Sompel: "The bX Project: Federating and Mining Usage Logs from Linking Servers". CNI Coalition for Networked Information, Dec. 5-6, 2005, Fall Task Force Meeting, Phoenix, Arizona. http://library.lanl.gov/cgi-bin/getfile?LA-UR-05-9439.pdf

Bollen J., M.A. Rodriguez and H. Van de Sompel : "Journal Status". *Scientometrics*, 69(2006)3, 669-687.

Bollen J., H. Van de Sompel, J.A. Smith and R. Luce: "Toward alternative metrics of journal impact: A comparison of download and citation data". *Information Processing & Management*, 41(2005)6, 1419-1440.

Butler, D.: "Academics strike back at spurious rankings". *Nature*, (2007)447, 514-515.

Creswell, J.:*Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks, California : Sage Publications, 2003.

Dong, P., M. Loh and A. Mondry: "The 'impact factor' revisited". *Biomedical Digital Libraries, 2*(2005)1, http://dx.doi.org/10.1186/1742-5581-2-7

Freeman, L.: *The Development of Social Network Analysis*. Vancouver : Empirical Press, 2004

Harnad, S.: "Open Access Scientometrics and the UK Research Assessment Exercise".*11th Annual Meeting of the International Society for Scientometrics and Informetrics*, Madrid, 2007. http://arxiv.org/abs/cs/0703131

KE: *Usage Statistics Report*. Copenhagen : Knowledge exchange, 2007. http://www.knowledge-exchange.info/Admin/Public/DWSDownload.aspx?File=%2fFiles%2fFiler%2fdownloads%2fIR+workshop+1617+Jan+2007%2fNew+reports%2fKE_IR_strand_report_Usage_Statistics_Sept_07.pdf

Kleinberg, J.: "Authoritative sources in a hyperlinked environment". *Proc. 9th Ann. ACM-SIAM Symp. Discrete Algorithms*, (1998), pp. 668-677.

Lawrence S., C.L. Giles and K. Bollacker: "Digital libraries and autonomous citation indexing". *IEEE Computer* 32(1999)6, 67-71

Monastersky, R.: "The Number That's Devouring Science". *The Chronicle of Higher Education*, (October 2005). http://chronicle.com/free/v52/i08/08a01201.htm

Scholze, F. and S. Dobratz: "International Workshop on Institutional Repositories and Enhanced and Alternative Metrics of Publication Impact, 20–21 February 2006, Humboldt University Berlin, Report". *High Energy Physics Libraries Webzine*, (October 2006)13. http://library.cern.ch/HEPLW/13/papers/2/

Seglen, P.O.: "Why the impact factor of journals should not be used for evaluating research", *BMJ, (1997)*314(7079), 498-502. http://www.bmj.com/cgi/content/full/314/7079/497

Shepherd, P.T.: *Final Report on the Investigation into the Feasibility of Developing and Implementing Journal Usage Factors*. United Kingdom Serials Group, May 2007. http://www.uksg.org/usagefactors/final

WEBSITES REFERRED TO IN THE TEXT

CNI - Coalition for Networked Information. http://www.cni.org/

COUNTER - Counting Online Usage of NeTworked Electronic Resources http://www.projectcounter.org

COUNTER Codes of Practice. http://www.projectcounter.org/code_practice.html

DFG - Deutsche Forschungsgemeinschaft. http://www.dfg.de/en/index.html

DINI - Deutsche Initiative für Netzwerkinformation. http://www.dini.de/

DOAJ - Directory of Open Access Journals. http://www.doaj.org/

DRIVER - Digital Repository Infrastructure Vision for European Research. http://www.driver-repository.eu/

Eigenfactor.org. http://www.eigenfactor.org/

IFABC - International Federation of Audit Bureaux of Circulations. http://www.ifabc.org/standards.htm

Journalprices.com. http://www.journalprices.com/

Knowledge Exchange. http://www.knowledge-exchange.info/

LIBER Access Division. http://www.libereurope.eu/node/15

MESUR - MEtrics from Scholarly Usage of Resources. http://www.mesur.org

METIS – Meldesystem für Texte auf Internet Seiten. http://www.vgwort.de/metis.php

OAI-PMH - Open Archives Initiative Protocol for Metadata Harvesting. http://www.openarchives.org/OAI/openarchivesprotocol.html

OpenDOAR - Directory of Open Access Repositories. http://www.opendoar.org

OpenDOAR policies tool. http://www.opendoar.org/tools/en/policies.php

SCI - Science Citation Index. http://scientific.thomson.com/products/sci/

SUSHI - Standardized Usage Statistics Harvesting Initiative. http://www.niso.org/committees/SUSHI/SUSHI_comm.html

UKSG – UK Serials Group. http://www.uksg.org/

NOTES

---

[1] From Wikipedia: http://en.wikipedia.org/wiki/Page_rank
[2] Journals are indexed for the Science Citation Index Expanded (November 2007).