

European Network of Indicator Designers (ENID) Conference in Rome, 7th-9th September 2011
Conference session: Coverage and completeness of research output in data sources

Janne Pölönen (Federation of Finnish Learned Societies, Finland), Yrjö Leino (CSC – IT Center for Science, Finland) & Otto Auranen (Federation of Finnish Learned Societies, Finland):

COVERAGE AND RANKING OF JOURNALS: COMPARISON OF SIX DATA SOURCES

Research subject

The coverage of publication and citation databases, especially the much used Web of Science databases, has been analyzed in several bibliometric studies (e.g. Bakkalbasi et al. 2006, Gavel & Iselid 2008, van Leeuwen et al. 2001, Neuhaus et al. 2006). Bibliometric methods are used for research evaluation, analysis of scientific publishing and in university rankings, which are increasingly common as tools of science policy. Hence the coverage of databases is a relevant issue both for bibliometric research community and science policy makers.

The coverage of publication databases is especially relevant in regard to publication output of social sciences and humanities. It is well known that social scientists and humanities scholars publish much of their research as monographs and articles in edited books, and in other languages than English. These publications are less extensively covered by for example Web of Science or Scopus databases. (Archambault & Gagné 2004, Hicks 1999, Hicks 2004, Nederhof 2006). The requirements of research evaluation have generated recent interest in the coverage of data sources of social sciences and humanities literature (Hicks & Wang 2010, 2011, Dassa et al. 2010).

This paper compares the coverage and quality indicators of six data sources for scientific journals in all fields of science. The study is part of the Finnish Publication Forum Project that aims at quality classification of scientific publication channels (journals, series and publishers), following the Norwegian model (Schneider 2009), for the performance based allocation of university funding. The project is funded by the Ministry of Education and Culture and based in the Federation of Finnish Learned Societies. The first part of the paper deals with the coverage of data and the second part with the journal rankings and their correlation with the impact factors.

Methodologies

The material of this study consists of journal data (titles, ISSN numbers, field classifications, quality rankings and impact factors) collected from six sources: the Thomson Reuters' Web of Science (including the *Science Citation Index*, the *Social Science Citation Index*, and the *Arts & Humanities Citation Index*), the title list of Elsevier's Scopus, the journal ranking lists produced in Norway, Denmark and Australia, and the European Science Foundation's *The European Reference Index for the Humanities* (ERIH).

Because of considerable overlap between the sources, an important part of the project has been to combine the journal data from all six sources into one list of unique journals. This has been done in collaboration with the CSC – IT Center for Science that removed the duplicates on basis of the journal titles and ISSN numbers. The result is a list of almost 36,000 unique journals in all fields of science.

Results

Our analysis confirms that Scopus has wider coverage of SSH journals than Web of Science, and that neither of the commercial databases covers this field as well as the national data sources of Norway, Denmark and Australia. Our analysis extends, however, also to the natural and medical sciences.

It should be possible to use the data of 36,000 journals to distinguish between indisputably international journals that appear in all or most sources, and journals of more national readership that appear only on one or two lists. Overall, we find that about one-fifth of all journals appear in all the sources, while around two-fifths of the journals appear only in one source, but there are field specific differences. Web of Science has the smallest proportion of journals that appear on no other list.

In the recent analyses, little attention has been paid to comparison of national journal rankings produced by peer evaluation, or to their correlation with the citation based impact factors. Indeed, the journal ranking data offers yet another avenue for comparison between peer-review and citation analysis (in general: Van Raan 2004, 20, Moed 2005, 229-245). Quality of journals is an important dimension of the journal coverage and the analysis of rankings can bring some interesting results: despite the weaker coverage of SSH journals in general, the Web of Science seems to have the best coverage of the ERIH level A journals.

There is an agreement over higher ranking publication channels in the three national lists in regard to around 1,000 journals that have been ranked on the level 2 or A (including A* in Australia) in all of them. Overall, it is found that an important share of journals given high rank in one national evaluation, or in the *European Reference Index for the Humanities*, has been considered leading also in the other rankings. We propose to investigate also the field specific differences.

The national rankings show some positive agreement with the citation impact factors: Thomson Reuters' *Journal Impact factor* (JIF), and Elsevier's *Source Normalized Impact per Paper* (SNIP) and *Scimago Journal Rank* (SJR). It is a constant pattern that in each ranking the group of top level journals has a higher average value than lower level journals. This is unsurprising in case of the natural and medical sciences, but we analyze also the correlations between citation based indicators and the ERIH ranking of humanities journals.

References

Archambault, E. & Gagné, V.G. (2004) *The Use of Bibliometrics in the Social Sciences and Humanities. Science-Metrix Final Report prepared for the Social Sciences and Humanities Research Council of Canada, Montreal.*

Bakkalbasi, N., Bauer, K., Glover, J. & Wang, L. (2006) "Three options for citation tracking: Google Scholar, Scopus and Web of Science", *Biomedical Digital Libraries*, vol. 3, no. 7.

Dassa, M., Kosmopoulos, C. & Pumain, D. (2010) "JournalBase - A Comparative International Study of Scientific Journal Databases in the Social Sciences and the Humanities (SSH)", *Cybergeog: European Journal of Geography*, January 2010.

Gavel, Y. & Iselid, L. (2007) "Web of Science and Scopus: a journal title overlap study", *Online Information Review*, vol. 32, no. 1, 8-21.

Hicks, D. (2004) "The Four Literatures of Social Science", in H. Moed, W. Glänzel & U. Schmoch (eds.), *Handbook of Quantitative Science and Technology Research*, Kluwer Academic Publishers, Dordrecht, 473-496.

Hicks, D. (1999) "The difficulty of achieving full coverage of international social science literature and the bibliometric consequences", *Scientometrics*, vol. 44, no. 2, 193-215.

Hicks, D. & Wang, J. (2011) "Coverage and overlap of the new social science and humanities journal lists", *Journal of the American Society for Information Science and Technology*, vol. 62, 284–294.

Hicks, D. & Wang, J. (2010) Towards a Bibliometric Database for the Social Sciences and Humanities – A European Scoping Project. A report produced for DFG, ESRC, AHRC, NWO, ANR and ESF, March 2010.
http://www.vandenbesselaar.net/_pdf/2010%20ESF.pdf

Moed, H.F. (2005) *Citation Analysis in Research Evaluation. Information Science and Knowledge Management*, vol. 9, Springer, Dordrecht.

Nederhof, A.J. (2006) "Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A Review", *Scientometrics*, vol. 66, no. 1, 81-100.

Neuhaus, C., Neuhaus, E., Asher, A. and Wrede, C. (2006) "The depth and breadth of Google Scholar: an empirical study", *Portal: Libraries and the Academy*, vol. 6, no. 2, 127-141.

Schneider, J.W. (2008) "An outline of the bibliometric indicator used for performance-based funding of research institutions in Norway", *European Political Science*, vol. 8, 364-378.

van Leeuwen, Th.N., Moed, H.F., Tijssen, R.J.W., Visser, M.S. & van Raan, A.F.J. (2001) "Language biases in the coverage of the *Science Citation Index* and its consequences for international comparisons of national research performance", *Scientometrics*, vol. 51, no. 1, 335-346.

Van Raan, A.J. (2005) "Measuring Science", in H. Moed, W. Glänzel & U. Schmoch (eds.), *Handbook of Quantitative Science and Technology Research*, Kluwer Academic Publishers, Dordrecht, 19-50.