How to evaluate academic research

A comment after the Thomson Reuters–UQ conference on the Perspectives on Metrics-Based Research Evaluation, Brisbane 16–17 April 2009.

Conducting research, along with teaching, is the most important function of a university. In Australia, funding for academic research is made available through federal government, allocated to researchers in the form of contestable research grants from one of the Australian research councils (ARC or NHRMC). The remainder comes in the form of a block grant allocated directly to universities by government based on a set of pre-defined criteria such as annual counts of research publications. Increasingly, funding is also sought from private enterprise (predominantly for joint R&D activities) as well as the charitable sector (mainly for medical research). In the 2006/7 budget the Australian government allocated 2 234.3 million dollars[1] to support research in Australia.

The government, as the main funder of university-based research, is increasingly interested in assessing the quality or impact of the research in order to inform its funding allocation, to collect information on areas of current research activity, to identify centres of excellence and areas needing additional investment, and to inform its science and innovation policy. To support this assessment, mechanisms by which the quality or impact of scientific endeavor can be measured are being developed, tested, and implemented. Traditionally, research quality has been evaluated by the process of peer review. However, this approach is proving too costly for large, national-based evaluation exercises, and alternative or supplementary approaches are required. One such approach, gaining support of both funders and researchers, is an approach based on the use of objective, quantitative indicators, including these based on publication and citation counts (bibliometrics).
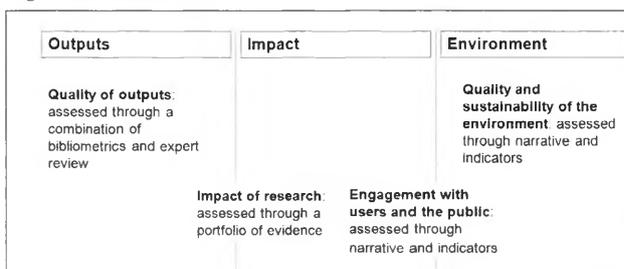
In Australia, the Excellence in Research in Australia framework (ERA) is the world's first research assessment exercise using metrics (predominantly bibliometrics) to evaluate and, in the future, fund research in Australia. Previous national schemes, such as those in the United Kingdom and New Zealand have, until now, based their assessments on peer review where portfoilios of publications and other evidence are reviewed by academic experts. The UK's new Research Excellence Framework (REF) and similar systems in Hong Kong and Singapore are now considering an increased use of metrics-based evaluation.

Recognising the importance of this shift, Thomson Reuters, the world-leading provider of citation data, joined forces with the University of Queensland to organise a two-day conference on national and institutional perspectives on metrics-based research evaluation. The conference attracted over 140 delegates from Australia and New Zealand, representing both research management and academic library communities from universities, research institutes, and funders of research. The main aim of the conference was to bring these two communities together to explore issues around supporting effective research evaluation as well as to share up-to-date research findings and current practice in the field.

The conference opened with a keynote address from Professor Charles Oppenheim, Head of the Department of Information Science at Loughborough University who spoke of the history and development of the UK's approach to evaluating research. The UK system, known as the Research Assessment Exercise (RAE) is the oldest national scheme, having received the first submissions in 1986, and was predominantly based on a peer review approach. Oppenheim discussed a recent shift from the RAE to the new REF, to be implemented in 2011 and based heavily on metrics-based quantitative data. He stated that this shift was mainly a political decision motivated by the UK government's desire to limit the costs of the research evaluation process (the belief that metrics can be gathered and compiled at the fraction of the cost of the peer review process). The decision was also supported by extensive research demonstrating correlations between rankings obtained through peer review exercises (e.g. RAE) and bibliometrics. Interestingly, Oppenheim, who is the government's expert advisor on the use of metrics in the REF, indicated that, despite the initial desire to shift to a predominantly metrics-based system, this is now not likely. The REF aims, as determined by the policy makers and through the sector consultation process, are now too broad to rely on a suite of pre-defined indicators (see fig. 1) and human intervention is required at least at the level of analysing and interpreting the data and for some aspects of the evaluation process – a detailed peer review.

Fig. 1 The REF Framework

Outputs	Impact	Environment
Quality of outputs: assessed through a combination of bibliometrics and expert review		Quality and sustainability of the environment: assessed through narrative and indicators
	Impact of research: assessed through a portfolio of evidence	Engagement with users and the public: assessed through narrative and indicators

Charles Oppenheim, 2009

A further keynote address was delivered by Professor Robert Tijssen from Centre for Science and Technology Studies (CWTS) at Leiden University in the Netherlands. His excellent talk concentrated on aspects of creating indicators-based assessment of the impact of research outside the academic sphere – especially mid- to long-term economic and social benefits. This aspect of evaluation is particularly difficult, and not often discussed by the bibliometrics research community, as the time frames of possible impact are much longer than in the academic domain (which is normally between 3 to 5 years) and the indicators are not as easily quantifiable or accessible to institutions or evaluators (Fig. 2). Nonetheless, Prof Tijssen suggested a number of indicators of impact (mid-term) and outcomes (long-term) and presented a comprehensive framework of indicators by which a research unit can be practically evaluated for impact.

Fig. 2 Outputs, Impacts, and Outcomes of research
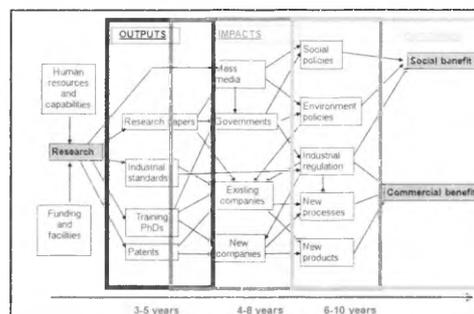
Robert Tijssen, 2009 (Adapted from Lewison, 2009)

The Australian approach to bibliometrics research evaluation was described by Andrew Calder from the Australian Research Council (ARC), the agency responsible for the design and implementation of the ERA. His presentation was more focused in scope and concentrated on the process of developing bibliometrics indicators and the description of the selected indicators. The Australian system, though relying heavily on indicators (at least in disciplines where these can be used reliably, i.e. physical sciences, biological sciences and medicine, and some social sciences), will still rely upon the advice of a panel of experts who will interpret available metrics. The bibliometric indicators for the assessment will be

supplied by participating institutions and will include lists of all research publications by ERA eligible staff. The journal publication data will be matched to journal lists collated by the ARC to calculate percentages of institution's publications in quality groupings – from A* (internationally excellent) to 'not rated'. Also, the numbers of citations to these publications will be harvested from citation indices and compared to 'world' and 'Australian' averages. In addition, a table showing the distribution of the institution's papers across percentiles (top 1st, 5th, 10th, 25th, and 50th percentiles) for each research area will be compiled.

Linda Butler, a visiting researcher at ANU and an adviser to the ARC, described reasons for creating Australian-specific journal classification and benchmarks and presented a study in which she analysed publication data based on the Australia and New Zealand Fields of Research-derived benchmarks and standard benchmarks based on Thomson Reuters journal classification and demonstrated the differences in institutional scores depending on the used classification scheme.

The approaches to research evaluation and the use of bibliometrics in that process in New Zealand were eloquently described by Warren Smart from the New Zealand Ministry of Education.

There were also two papers addressing the issue of university rankings. Prof Liu from Jiao Tong Shanghai University described the development of probably the most famous university ranking system in the world while Tony Sheil from Griffith University spoke of possible approaches to ranking smaller universities, which often cannot be adequately assessed through global listings, to identify their areas of strength and benchmark against institutions with similar areas of research activity.

The afternoon session of Day 1 had very interesting papers describing how institutions use bibliometrics to drive academics' behaviour and improve the quality of their published outputs. Professors Owens and Lopez described the approaches at the University of Western Australia (SOCRATES index), and School of Population Health at UQ respectively. Prof Lopez, who has been the Head of School for the last seven years, was able to demonstrate how setting performance standards and rewarding achievement improved the quality of research at his school as demonstrated by increased publications in international, high impact journals and increased success rate in grant applications. Prof Owens described the principles of the SOCRATES index – a score that each UWA academic is given based on a number of performance indicators. One of the elements of the score is the numbers of publications and citations as recorded in Thomson Reuters' Web of Science. The interesting discussion point was the technical approaches in which external data (like WoS publication and citation counts) can be integrated with internal university systems.

This presentation nicely directed the discussion towards the main theme of Day 2 of the conference when the role of the academic library in supporting research management and evaluation were discussed. We began with a keynote by Keith Webster, University Librarian and Director of Learning Services at the University of Queensland. Webster spoke about the changing role of both university library and university librarian and the perfect positioning of the library, library technology, and professional skills set of librarians to play an increasingly important role in the management of research and research management data and support of research evaluation process. The following three papers outlined practical examples of university libraries supporting the research evaluation process – UNSW, directly by providing sophisticated bibliometric analyses of school, research centres, and individual academics, UQ by integrating publication collection (so called HERDC collection) process within its institutional repository and creating definitive publication lists for individuals and academic

units augmented by citation information and UNISA by supporting publication collection for ERA. A recurring theme through these presentations was the difficulty of compiling reliable publication lists for individual researchers and the need for creation and implementation of unique author identifier. Two papers addressed that issue: The Australian National Library's People Australia and Thomson Reuters' ResearchID projects were described by Basil Dewhurst and Ellen Rotenberg respectively.

Finally, the conference ended with a 1/2 day masterclass on the principles of evaluating research excellence. Professor Robert Tijssen led the session in which he defined the concept of "research excellence" and proposed a set of indicators which could be used to assess the research performance.

The conference was very well received by both delegates and invited speakers. In the words of one delegate: "I found the conference to be of a uniformly high standard, with every presentation interesting, informative, and stimulating. The keynote addresses provided a wonderful overview of developments worldwide, while the local talks gave some fascinating insights into the way in which various universities, across their research and library portfolios, are responding to both the ERA and global university rankings. The collaboration of Thomson Reuters with the university sector on the development of robust metrics and visualisation tools for research performance is calibrating the global research community to the highest standards, and will lead to better research and better research outcomes for the communities we serve."

Information on program, speakers, and their presentations are available to delegates at **http://scientific.thomsonreuters.com/ ausbiblioconference**

Dr Berenika M. Webster
Thomson Reuters
berenika.webster@thomsonreuters.com