Institutional repositories and research assessment

Michael Day[1]

UKOLN, University of Bath, Bath BA2 7AY, United Kingdom
m.day@ukoln.ac.uk

Abstract. This study concerns the potential role of institutional repositories in supporting research assessment in universities with specific reference to the Research Assessment Exercises in the UK. After a brief look at research evaluation methods, it introduces the UK Research Assessment Exercise (RAE), focusing on its role in determining the distribution of research funding, the assessment process itself, and some concerns that have been raised by participants and observers. The study will then introduce institutional repositories and consider the ways in which they might be used to enhance the research assessment process in the UK. It will first consider the role of repositories in providing institutional support for the submission and review process. Secondly, the paper will consider the ways in which citation linking between papers in repositories might be used as the basis for generating quantitative data on research impact that could be used for assessment. Thirdly, this study will consider other ways in which repositories might be able to provide quantitative data, e.g. usage statistics or Webometric link data, which may be able to be used - together with other indicators - to support the evaluation of research.

1. Introduction

The ePrints UK (http://www.rdn.ac.uk/projects/eprints-uk/) project is funded by the Joint Information Systems Committee (JISC) as part of the Focus on Access to Institutional Resources (FAIR) Programme. The aim of the project is to develop a national service that gives access to e-print records, derived by harvesting metadata from institutional and subject-based repositories using the Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH). In addition, the project aims to provide access to these institutional assets through Resource Discovery Network (RDN) faculty level hubs and the Education Portal. It is also investigating the use of Web Services technologies for the enhancement of metadata and for the automatic linking of citations.

The first ePrints UK supporting study introduced the project and assessed the prospects for institutional repositories in the UK (Day, 2003). Other studies are looking at collection development and the business and intellectual property rights issues surrounding the sharing and enhancement of metadata. This study concerns the potential role of institutional repositories in supporting research assessment in universities with specific reference to the Research Assessment Exercises in the UK.

1 A supporting study for the ePrints UK project, v. 0.1 (draft), 2 December 2004

The opening section of this study will provide a brief assessment of methods of research assessment focusing on the use of peer-review and quantitative indicators. The following section will focus on the UK Research Assessment Exercise (RAE), and in particular its importance in determining the distribution of research funding, looking at the assessment process itself, the results of the 2003 Roberts review and plans for the next RAE in 2008; highlighting a number of issues that have been raised by participants and observers. The study will then introduce institutional repositories and consider in more detail the ways in which they might be used to enhance the research assessment process in the UK. It will first consider the role of repositories in providing institutional support for the submission and review process. This is an approach that would build on existing approaches to the RAE within institutions and might possibly help provide a partial rationale for the creation of institutional repositories, at least in the short-term. Secondly, the paper will consider the ways in which citation linking between papers in repositories might be used as the basis for generating quantitative data on research impact that could be used for assessment. Some work has already been done on this as part of the JISC and US National Science Foundation funded Open Citation Project (Hitchcock, et al., 2002). The ePrints UK project is itself investigating the use of Web Services technologies for facilitating the automatic citation linking of e-prints. In addition, Harnad, et al. (2003) have developed a proposal for research active staff in UK universities to maintain standardised *curriculum vitae* linked to the full-text of refereed research papers that could be used for "online harvesting, scientometric analysis and assessment." Thirdly, this study will consider other ways in which repositories might be able to provide quantitative data, e.g. usage statistics or Webometric link data, that may be able to be used - in combination with other indicators - to support research assessment.

2. Methods of research evaluation

In recent years, research evaluation has become an important issue for many universities and research institutes. Geuna and Martin (2003, p. 277) attribute this to the twin demands of accountability and the need to demonstrate the efficient use of public expenditure on higher education and research. The UK RAEs reflect a growing trend for governments to set up research evaluation exercises as a means of determining the allocation of resources. Evaluation can take place at several different levels, e.g. that of individual researchers, research groups, university departments or institutions. In practice, however, national evaluation exercises tend to focus on the group or department as the main 'unit of assessment,' although evaluations at lower levels of granularity could be used to inform this. The main evaluation methods used include peer review and the application of quantitative methods.

2.1 Peer review

In the context of national research evaluation exercises, peer review usually means that a committee of scholarly peers reviews those research units being assessed in

accordance with some pre-defined criteria. One of the advantages of peer review over other approaches is its widespread use elsewhere in the academic world, e.g. as part of the publication process and for deciding the allocation of research grants. With regard to publishing, peer review has been described as the "principal social mechanism for quality control in academic science" (Ziman, 2000, p. 42).[2] The use of peer review for the assessment of grant applications is more controversial. Horrobin (1996) notes that the rejection of a research proposal "may well ensure that the research in question will never be done at all." While peer-review has sometimes been criticised for its lack of fairness, lack of reliability, and inherent conservatism, few have argued that it should be completely phased out (Wessely, 1998). In national research evaluation exercises, peer review is often supplemented by data on research funding, publications or citations, a method that Geuna and Martin (2003, p. 279) call "informed peer review."

2.2 Quantitative methods

The types of metrics used for research evaluation typically include both input and output measures. Input measures might include the amount of research funding awarded or staff numbers, factors that are taken account of in the UK RAEs. Output measures are typically based on analyses of publication counts, citations or the number of patents issued (King, 1987). Proponents of the use of quantitative indicators for research evaluation argue that they are considerably cheaper than peer-review and potentially more objective (e.g., Oppenheim, 1996; Williams, 1998).

Perhaps the most widely used output measures used for research evaluation are those based on bibliometric analysis, chiefly the statistical analysis of citation data. Bibliometrics had its origins in the development of citation indexes for bibliographic searching towards the end of the 1950s. The publication of the *Science Citation Index*, first produced by the Institute of Scientific Information in 1961, enabled sociologists of science for the first time to investigate the "structure of science and the process of scientific development" (Garfield, 1983, p. 62). By the 1970s, however, there was a growing interest in how citation analysis could be used to support the evaluation of research, even where this might entail the redistribution of scientific resources (Wade, 1975).

Until recently, the main sources of bibliometric data have remained the citation indexes produced by the Institute of Scientific Information (now Thomson ISI). These indexes capture descriptive metadata for articles published in selected journals and manually link all citations into a database that can be used for bibliographic searching and various types of bibliometric analysis. For example, Thomson ISI themselves analyse this data to produce some basic citation measures. The annual Journal Citation Reports service provides data on over 7,500 scholarly and scientific journals, including the number of papers published, citation counts, and journal impact factors (http://www.isinet.com/products/evaltools/jcr/). Controversially, some of this data - in particular the journal impact factor - is sometimes used as a surrogate measure of research quality.

---

2 Despite this, a systematic review of the use of peer-review by biomedical journals has found few well-designed studies assessing its effects (Jefferson, et al., 2002).

Impact factors were first devised by the Institute of Scientific Information as a means of evaluating journals and are intended to be used in a variety of ways, e.g. to support library collection development decisions, to help authors with decisions on which journals they should submit papers to, and for the study of scientific trends through bibliometric analysis (Glänzel & Moed, 2002). As defined by Thomson ISI's Journal Citation Reports, impact factors are ratios calculated by dividing the citations received by a journal in a given year by the number of papers published in that journal over the previous two years (Garfield, 1996). To give a simple example, if in 2001 a journal has received 50 citations to articles published in 1999-2000 and the total number of articles published by that journal in 1999-2000 was 150, the impact factor for 2001 would be 0.333. If over the same period of time it received 70 citations, the impact factor would be 0.466. Journal impact factors enable the comparison of journals while taking account of variability in the number of papers published each year and other factors.

ISI have always been careful to stress that bibliometric data like impact factors should only be used to supplement qualitative measures in the evaluation of research groups or individuals. However, there is evidence that impact factors have been used in countries like Italy and Japan to drive recruitment or research funding decisions (e.g., Jennings, 1998). To take an extreme case, a news feature in *Nature* highlighted how funding for university hospitals in Finland partly depends on publication points, "with a sliding scale corresponding to the impact factor of the journals in which researchers publish their work" (Adam, 2002). In Germany also, the research funding of university departments and the appointment of science faculty is often based on the impact factors of the journals in which they publish (Oehm & Lindner, 2002; Kaltenborn & Kuhn, 2003), although the Deutsche Forchungsgemeinschaft (1998) has stressed that an 'adequate' evaluation of individuals or small groups should be based on qualitative criteria, i.e. that "their publications must be read and critically compared to the relevant state of the art and to the contributions of other individuals and working groups."

This use of journal impact factors for the evaluation of individual researchers or groups is extremely contentious and has generated a substantial literature, especially in the biosciences (e.g., Opthof, 1997; Hecht, Hecht & Sandberg, 1998; Gisvold, 1999). Moed (2002) warns that bibliometric indicators like impact factors reflect impact, not quality, arguing that they can only be used as supplementary tools in the evaluation of research, and only then if their inadequacies are taken into account. Critics like Seglen (1997) have pointed out that the citation rates for individual papers do not correlate with the impact factors of the journal in which they are published. Citation rates tend to be skewed, meaning that heavily cited articles distort the impact factor of the journals (Seglen, 1992). The ISI dataset itself has also been subject to some criticism. For example, van Leeuwen, et al. (2001) found that the inclusion of non-English language journals in the Science Citation Index 'diluted' the impact score of countries like Germany, France and Switzerland. The implications of this are that ISI's data should be used with care for international comparisons of research performance, e.g. those undertaken by May (1997, 1998), Adam (1998) and King (2004), or the production of university rankings like those produced by Shanghai Jiao Tong University (http://ed.sjtu.edu.cn/rank/2004/2004Main.htm).[3]

[3] Van Raan (2004) has provided a recent critique of these university rankings.

After uncovering some inaccuracies in the citation statistics compiled by ISI, the editors of *Nature* (2002) recommended that researchers, policy makers and publishers should treat them with caution.

3. The UK Research Assessment Exercises

The Research Assessment Exercises are periodic evaluations of the research undertaken by UK universities initiated by the UK higher education funding bodies[4] in order to inform the allocation of research funding. The University Grants Committee (UGC) introduced the first 'research selectivity exercise' in 1985-86, with further evaluations being performed in 1989, 1992, 1996, and 2001. The next exercise is scheduled for 2008 (http://www.rae.ac.uk/).

3.1 The RAEs and the allocation of research funding

Public funding for research in universities and other higher education institutions in the UK is provided by a system of 'dual support.' Firstly, the research infrastructure is funded by the education departments through the higher education funding bodies in the form of a 'block grant' paid to each institution. This funding is intended to support a basic level of research activity and covers, for example, the salaries of permanent members of staff, the cost of premises and central facilities (libraries, computing services, etc.) and can be spent entirely at the institution's discretion. The second stream of public funding comes from the Department of Trade and Industry's Office of Science and Technology (OST) and is distributed through the research councils. This is used to fund specific research projects and some central facilities. Additional project based research funding is also available from a wide range of other sources, including charities, government departments, industry, and through initiatives like the European Union's Framework Programmes.

From the start, the UGC saw the Research Assessment Exercises as a basis for informing the selective allocation of the block grant paid to higher education institutions for research. In 1984, the committee had already recommended the adoption of a "more selective approach in the allocation of research support among universities in order to ensure that resources for research are used to the best advantage" (UGC, 1984, 1.9). The proportion of the available funds allocated on the basis of the exercises has risen steadily since the 1986 exercise. Since 1992, almost all of the funding allocation for research from the funding bodies has been based on the results of the RAE (Stiles, 2002).

The exact formula used to allocate funding varies between the four funding bodies, but in general, it is based on quality (the RAE rating awarded) and volume (typically the number of research active persons submitted to the RAE), weighted to benefit those units awarded the highest ratings. The formulae used to allocate funding

[4] These are the Higher Education Funding Council for England (HEFCE), the Scottish Higher Education Funding Council (SHEFC), the Higher Education Funding Council for Wales (HEFCW), and the Department for Employment and Learning Northern Ireland (DELNI)

have evolved over the lifetime of the exercises, progressively limiting funding to the best-rated units of assessment within each institution. For example, after the 1992 exercise, some research funding was allocated to all units except for those that received the lowest RAE rating of 1 (Ball, 1997, p. 284). Following the most recent exercise (RAE 2001) the English funding council only allocated research funding to units with a RAE rating of 4 or above (HEFCE, 2003). The Higher Education White Paper of 2003 proposed an even more extreme level of selectivity, arguing that research concentration "brings real benefits, including better infrastructure (funding excellent equipment and good libraries), better opportunities for interdisciplinary research, and the benefits for both staff and students that flow from discussing their research and collaborating in projects" (DfES, 2003a, 2.8). In the short term, the White Paper suggested that HEFCE should use the results of RAE 2001 to help identify the 'very best' 5* rated departments in order to direct more funding to these in advance of RAE 2008 (DfES, 2003a, 2.15). In order to fund this extra funding, the Department for Education and Skills suggested that the funding for units rated 4 should be frozen until the next exercise (DfES, 2003b). With increases in funding available, however, HEFCE's actual response to this in its grant allocations was less drastic (Goddard, 2004; HEFCE, 2004).

The general tendency of the exercises, when combined with the changing funding formulae adopted by the funding bodies, has been to widen the gap between the highest and lowest ranking departments (Ball & Butler, 2004 p. 91).

3.2 The RAE process and methodology

At the time of writing, the evaluation methodology of the RAE is broadly based on a peer-review process. In 2001, for example, those organisations being assessed - typically university departments, faculties or schools - prepared a standardised submission that would then be reviewed by expert panels in one of 68 discipline-based units of assessment. The submission would have included information on staff, the number of research students, grant income, research strategy, infrastructure and 'evidence of esteem.' In the 2001 RAE, each 'research-active' member of staff included in the evaluation process could submit up to four research outputs (chiefly publications) for review by the assessment panel.

For RAE 2001, panel chairs were nominated by members of the 1996 RAE panels and appointed by the four funding bodies. Panel members are nominated by a wide range of organisations, but appointed by the funding bodies on consultation with the panel chairs (ref). The size of panels varied, but their membership largely comprised UK academics with additional representatives from learned societies, professional associations, government departments and other organisations. For the 2001 exercise, assessment panels were able to agree their precise criteria and working methods within a standardised framework developed by the funding bodies. For example, they could determine the exact weight given to published outputs and other indicators of peer esteem (Bessant, et al., 2003, pp. 52-53). For transparency, criteria and working methods for all panels were published (RAE Circular 5/99) over a year before submissions were due. For the 2001 exercise, panels could 'cross-refer' parts of submissions or selected outputs to other panels, to help deal with those areas where

disciplines overlap. Submissions and outputs are then considered by the panel and rated. In 2001, this was on a seven-point scale from 1 to 5* (Table 1).

Table 1: The rating scale from RAE 2001

5* (five star)	Quality that equates to attainable levels of international excellence in more than half of the research activity submitted and attainable levels of national excellence in the remainder.
5	Quality that equates to attainable levels of international excellence in up to half of the research activity submitted and to attainable levels of national excellence in virtually all of the remainder.
4	Quality that equates to attainable levels of national excellence in virtually all of the research activity submitted, showing some evidence of international excellence.
3a	Quality that equates to attainable levels of national excellence in over two-thirds of the research activity submitted, possibly showing evidence of international excellence.
3b	Quality that equates to attainable levels of national excellence in more than half of the research activity submitted.
2	Quality that equates to attainable levels of national excellence in up to half of the research activity submitted.
1	Quality that equates to attainable levels of national excellence in none, or virtually none, of the research activity submitted.

Source: RAE 2/99.

3.3 The impact of the RAEs on national research productivity

Since their introduction, the research assessment exercises have gradually become a familiar - if still not uncontroversial - part of academic life in UK higher education institutions.

Bibliometric analysis suggests that the RAEs have (at least) coincided with an increase in the international impact of UK research. Separate 'benchmarking' studies funded by the Office of Science and Technology (OST) and HEFCE in the mid 1990s used bibliometric data to compare the research performance of the UK with other countries. The initial OST study (May, 1997) found that the UK's percentage share of publications and citations between 1981 and 1994 was second only to that of the United States, and that it was in the top ten countries when measured by relative citation impact. The UK performed particularly well compared with other G7 countries when scientific output (in terms of citations) was measured in relation to government money spent on research and development. Follow-up research has shown the continued international importance of UK research (Evidence UK, 2003; King, 2004). The HEFCE study (Adams, *et al*., 1997; Adams, 1998) provided a comparative analysis of English research grouped by the units of assessment used in the 1996 RAE. The study, intended to identify strengths and weaknesses in research performance, concluded that England ranked 1st or 2nd in terms of research impact (citations per paper) in around half of the subjects analysed at unit of assessment level and that a high level of performance persisted when these units were clustered

with similar disciplines (Adams, et al., 1997, p. 10). Adams (1998) noted particular strengths in the core biological and physical sciences units of assessment. Adams (2002) has argued that the incentives associated with the RAEs led to measurable changes in research efficiency; saying that the exercises created incentives that channelled natural research competitiveness into a "pervasive driver of excellence."

These increases in research performance may or may not be entirely attributed to the RAEs, although it would be difficult to argue that the exercises have had no effect at all (Nature, 1997). In practice, however, a rapid increase in the quality of research creates problems for the RAE and UK research funding more generally. Following the 2001 exercise, which saw a massive increase in the number of units rated at 5 or 5*, Adams (2002) commented that "with some 55% of UK academics in the top-rated category, the funding differentials are being flattened, and the resource pot hasn't grown as fast as the improvements." This is one reason why the UK funding bodies commissioned the Roberts review into the future of research assessment.

3.4 Criticism of the RAEs

The improvement in national research performance, as measured by these OST and HEFCE studies, has not meant, however, that the exercises have been free from all criticism. Critics fall into two main categories. Firstly those who have fundamental objections to research assessment in general, who note that the RAE has unintended consequences in terms of academic 'game-playing,' the deterrence of collaborative research, and a downgrading in the importance of teaching and the regional role of universities. Other critics focus more on the process and methodology adopted by the exercises.

3.4.1 Fundamental objections to assessment

Perhaps the most basic objection to the RAE is that because the assessments are used to allocate funding it falls foul of what economists refer to as Goodhart's Law. This, as restated by Strathern (1997), says that, "when a measure becomes a target, it ceases to be a good measure." So, for example, Elton (2003) argues that performance indicators, "when used for control, are unreliable: they do not measure performance itself, distort what is measured, influence practice towards what is measured and cause unmeasured parts to get neglected." Others have argued that the RAEs have led to universities downgrading the importance of teaching (Brinn, Jones & Pendlebury, 2001) and a neglect of their important regional role (Ball, 1997). Some of the effects of the RAEs *have* been unintended. For example, rules on the submission of staff in earlier exercises sometimes led, at least anecdotally, to a football-style 'transfer market' for those regarded as 'star' researchers (e.g., MacLeod, 2004, 9 November). Other critics object to the wider cultural effects of the exercises. Willmott (2003) views the RAEs as a means of more tightly coupling research activity with the perceived 'needs of industry' and legitimising dependence on commercial funding.

The financial importance of the RAE to higher education institutions has meant that over the years they have developed successful strategies for improving rankings. Some commentators have even questioned whether the improvement in RAE ratings between the 1996 and 2001 exercises reflect "a genuine improvement in quality or

simply an improvement in playing the game" (Ball & Butler, 2004). Strategies can range from the simple use of RAE 'dummy-runs' for the identification of weaknesses before the real exercise, the exclusion of selected staff from the submission process, to the use of RAE-centric criteria in the evaluation of individual staff members. For example, it was reported in 2004 that academics in the medical school at Imperial College had been threatened with disciplinary procedures if "they did not generate external research grants of £75,000 a year and have at least three papers accepted for publication in peer-reviewed journals" (Fazackerley, 2004, 21 May). Some recent departmental 'restructurings' have been attributed to 'low' departmental RAE scores, e.g. the recent proposed closures of grade 4-rated departments at Exeter (chemistry), Newcastle-upon-Tyne (physics) and Cambridge (architecture). The increased concentration of research funding on the highest graded units seems to have encouraged university management to direct their focus only on those departments that have scored highly in former exercises or are likely to do so in future ones (Fazackerley, 2004, 26 November). In a what appears to be a demonstration of the Matthew effect (Merton, 1968) in action, research concentration seems to have had additional benefits for those institutions already strong in research. In his comparative analysis of national research productivity, King (2004) uses the Shanghai university rankings to conclude that the RAE has led to a concentration of research excellence in the UK centred on Cambridge, Imperial College, Oxford and University College London.

3.4.2 Criticisms of process and methodology

Other objections to the RAE have focused less on its wider effects but centre on aspects of process and methodology.

Transparency

The earliest research selectivity exercises were severely criticised for a lack of transparency (e.g., Smith, 1988; Phillimore 1989; Johnes & Taylor, 1992; Williams, 1998). In order to help address this issue, the funding bodies have for the most recent exercises published the criteria and working methods developed by the assessment panels (e.g., RAE Circular 5/99). In addition, a number of panels have published post-evaluation accounts of their deliberations. For RAE 2001, these include the assessment panels for business and management studies (Bessant, et al., 2003), town and country planning (Punter, et al., 2002), and library and information management (Elkin, 2002).

Cost

The Research Assessment Exercises are very expensive both in terms of direct costs and opportunity costs. For example, Lloyd has noted that the process places considerable demands upon and draws down on resources across the whole academic system (Punter, et al., 2002, p. 353). The Roberts review noted that the cost of the 1996 RAE (including opportunity costs) had been estimated as being between £27 million and £37 million, noting in mitigation, however, that this would represent just 0.8% of the total funds distributed on the basis of the exercise (Roberts, 2003, p. 21).

The neglect of bibliometric indicators

One of the most consistent criticisms of the RAE has been its seeming neglect of quantitative indicators. The use of such indicators (metrics) is seen as being considerably cheaper than peer-review processes and potentially more objective (Oppenheim, 1996; Williams, 1998). While RAE 2001 did take account of selected input measures like grant funding and the number of research active staff, decisions on the use of output measures like bibliometrics were delegated to individual panels.

There have been a variety of attempts made to correlate bibliometric indicators and RAE rankings, a task that has been made easier now that RAE submission information has been made available on the Web (http://www.hero.ac.uk/rae/). In a series of articles, Oppenheim and his colleagues at Loughborough University have studied the units of assessment for library and information science, genetics, anatomy and archaeology in the 1992 exercise (Oppenheim, 1995; 1997) and those for library and information management and archaeology in the 2001 RAE (Holmes & Oppenheim, 2001; Norris & Oppenheim, 2003). The studies showed a high statistically significant correlation between the RAE ranking and citation counts, findings also replicated in psychology by Smith and Eysenck (2002). Norris and Oppenheim (2003, p. 728) have concluded, that while citation analysis "is not a perfect tool, it is recommended that it should be adopted as the primary procedure for the initial ranking of university departments."

RAE panels have been free to use metrics as part of their evaluation strategy, although those that have published their deliberations stress the importance of reading all the outputs submitted to them (e.g. Punter, et al., 2002). However, an editorial in *Nature Neuroscience* (1998) noted a "widespread perception" that journal impact factors may "weigh heavily in many panels' recommendations."

The use of statistical methods to check influences on RAE

Other studies have used statistical methods to evaluate the ratings of the RAE and individual panels. Statistical analyses of the 1989 and 1992 exercises were used to identify some of the factors that influenced research performance and to test whether the data collected for the exercises could be used to construct useful indicators of research activity (Johnes, Taylor & Francis, 1993; Taylor, 1995). These studies found, for example, that there was a positive relationship between ratings and unit size and that the mean research rating is lower in units of assessment with a high proportion of 'new' universities. A detailed analysis of the unit of assessment for business and management studies in the 1992 RAE was used to explore the decision criteria used by the panel in their evaluation (Doyle, et al., 1996). In addition to Taylor's findings on size and 'new' universities, the study revealed that panel membership did appear to have a positive effect on rating (see also, Roberts, 1999), as did regional location. They also analysed the number of journal articles published in the 'top' British/European and US journals - as defined by ten-year impact factors (Doyle & Arthurs, 1995) - but found no evidence that these distinctions had any effect on the final ratings. Doyle, et al. (1996, p. 21) found this troubling, having assumed that publishing in the best US journals would be a good indicator of international excellence. In a more recent study, McKay looked at the unit of assessment for social policy and administration in the 2001 RAE and found that metrics on a relatively small number of criteria - size (the number of staff submitted), *per capita* research income, the number of doctorates awarded, publication in a 'core'

set of journals or of books by major publishers - could explain "over 80 per cent of the variation in actual ratings" (McKay, 2003, p. 457). He argued that statistical approaches, as well as being cheaper, might help to reduce the subjectivity inherent in qualitative assessment as well as the potential for bias.

None of these studies concluded that the use of quantitative methods could completely replace the peer-review process of the RAE. For example, Norris and Oppenheim (2003, p. 728) argued that peer review of written evidence should complement an initial ranking based on citation analysis. Doyle, et al. (1996, p. 25) proposed that the intelligent use of data analysis would liberate panels to concentrate their efforts "where its expertise and judgement have maximal value."

Possible bias

Another criticism of the research assessment exercises is its apparent bias. We have already noted how both Taylor (1995) and Doyle, et al. (1996) identified a positive relation between unit size and RAE rating. Other commentators have noted a perceived bias against applied research, consultancy, and clinical research in medicine [ref]. In response, the guidelines for RAE 2001 (RAE Circular 5/99) asked panels to "give full recognition to work of direct relevance to the needs of industry and commerce, and that all research, whether applied or basic/strategic, should be given equal weight."

Despite his own reservations, Tomlinson (2000) thinks that research assessment is here to stay, arguing that those who call for the abandonment of the RAE "will have to come up with a credible alternative for the accountable allocation of almost £1bn of public money."

3.4 The Roberts review and its implementation

The large increase in the number of units rated at 5 or 5* between the 1996 and 2001 RAEs promoted a period of reflection and review (Ball & Butler, 2004, p. 95). For example, the House of Commons Select Committee on Science and Technology held hearings on the RAE in January 2002 and again in May 2004. It published a report in April 2002 with a long list of recommendations and conclusions (House of Commons Select Committee on Science and Technology, 2002). While not all of the committee's recommendations were directly relevant to the RAE, the report noted the broadly beneficial effect of the RAE on UK research while expressing concern about its cost and frequency, the possible neglect of other university functions (e.g., teaching, community involvement, commercial activity), and the effects of the exercises on staff morale.

One underlying problem was that the dual support funding system itself was coming under strain due to imbalances in the money available from the research councils and the funding bodies. The pressure group Save British Science reported that the research councils' budgets had increased 123% (in real terms) since 1987, compared with a 26% increase in funding council budgets (Goddard & Davis, 2003). The resulting gap in funding had been estimated to be in excess of £170 million. This imbalance led to the OST proposing changes to the dual support system whereby higher education institutions would be required to identify and recover the full economic costs of research, although the research councils would only be obliged to

pay a proportion of these (OST, 2003). In response, some groups have concluded that the dual support system has now outlived its original purpose. For example, the Council for Industry and Higher Education argued that the UK should move to a position where the full costs of research are paid by research councils and other external funding bodies (Tysome, 2003). In November 2003, the Royal Society called for a fundamental review of university research funding, asking why grant income alone might not form the basis of calculating the infrastructure and indirect costs currently provided by the funding councils through the RAEs (Royal Society, 2003). An editorial in the *Times Higher Education Supplement* (2003) warned, however, that university staff should not be too sure that the RAE would disappear with the end of the dual support system, noting that it "is too useful to too many people not to continue in some form."

In the meantime, however, the four higher education funding bodies had invited Sir Gareth Roberts to review the research assessment system. His review solicited submissions from interested parties, held workshops and meetings, and funded an operational review of RAE 2001 (Farrant, Billing & Temple, 2003) and a study of international approaches to research assessment (Tunzelmann & Kraemer Mbula, 2003). The final report of the Roberts review was published for consultation in the summer of 2003 (http://www.ra-review.ac.uk/). It proposed the retention of a UK-wide system of research assessment dependent upon peer review. The review recommended an increase in the assessment period, to six years, with a mid-term monitoring of changes in the levels of research activity. It argued that the seven rankings used in the 2001 exercise should be replaced by a 'quality profile' of each submission that would allow for the production of a continuous grading scale, which would help eliminate some of the problems created by grade boundaries (Roberts, 2003, p. 10). Roberts also proposed that the relative value of these rankings, in terms of funding, should be published in advance of the exercise so that there was clear information available on the relationship between assessment scores and funding. The review proposed a reduction in the number of panels, suggesting between 20 and 25 unit of assessment panels supported by around 60 sub-panels.

One of the issues that the Roberts review had to address was the desirability, or otherwise, of giving performance indicators a greater role in research assessment. The use of performance indicators (e.g., those based on bibliometrics, grant income, etc.) has often been proposed as a cheaper and more objective alternative to peer-review, and they are used for research evaluation in some other countries (Guena & Martin, 2003; Tunzlemann & Kramer-Mbula, 2003). Out of the 414 responses to the initial invitation to contribute to the review, however, very few argued that metrics should be the *principal* means of assessing research. Rather, there was a widespread suspicion of metrics as "too crude a means to assess the quality of research (even in the hard sciences), and ill-suited to judge culture and the strategy and vision required to attain research excellence" (Roberts, 2003, p. 82). There was, however, support from a majority of respondents for the increased use of metrics in supporting the work of panels.

After a consultation period, the funding bodies announced their response to the Roberts review in early 2004 (HEFCE, et al., 2004). They announced that the next RAE would be held in 2008 and would still be based on peer-review by discipline-based panels. Following the recommendations of the Roberts review, the funding bodies said that the RAE would be held on a six-year cycle, that there would be 15 to

20 main panels supported by around 70 sub-panels, and that the rating scale would be replaced by a continuously graded quality profile with 4 starred levels. Particular attention would be directed towards reducing the administrative burden of the exercise. The Roberts review's proposal for mid-term monitoring, however, was not implemented.

4. Institutional repositories and research assessment

The UK research assessment exercises have evolved over the years to take account of the concerns of researchers and the changing nature of research in higher education itself. One of the most important of these changes has been a renewed emphasis on improving end-user access to the products of scientific and scholarly communication. Perhaps the most important catalyst of change was the advent of the World Wide Web in the 1990s. While experiments with the electronic publishing of journals began in previous decade, the Web provided an added impetus for publishers to make journals available online. By 2003, Cox and Cox (2003) could estimate that 75 per cent of all scholarly journals were available online.

At the same time, there has been an increased emphasis on the use of the Internet to make the results of science and scholarship freely available. For example, the Budapest Open Access Initiative (http://www.soros.org/openaccess/) advocates the removal of access barriers to the peer-reviewed journal literature, recommending both the 'self-archiving' by researchers of papers in open access repositories (e.g., Harnad, 2001) and the exploitation of 'open-access' business models for journal publishing. Examples of the latter approach include the growing list of 'author-pays' journals being published by BioMed Central (http://www.biomedcentral.com/) and the Public Library of Science (PLOS) (http://www.plos.org/). A number of research funding bodies have declared their support of open access principles, including the UK-based Wellcome Trust (2003), the French *Centre national de la recherche scientifique* (CNRS) and the *Max Planck Gesellschaft* (http://www.zim.mpg.de/openaccess-berlin/). In September 2004, the US National Institutes of Health (NIH) sought comments on a proposal to request that its grantees should be asked to provide the agency with an electronic copy of all published papers that would be made available in the National Library of Medicine's PubMed Central repository (http://www.nih.gov/about/publicaccess/). In the UK, a report of the House of Commons Select Committee on Science and Technology (2004a) recommended (amongst other things) that the research councils and other Government funding bodies should *mandate* their grantees to deposit a copy of published output in institutional repositories. While the UK Government's response to the report was generally lukewarm, e.g. viewing repository development as an issue for individual institutions to resolve (House of Commons Select Committee on Science and Technology, 2004c), Research Councils UK (RCUK) is currently consulting on a policy statement on publication of and access to research outputs that could make deposit in a repository a condition of being awarded a research council grant [ref.].

Various types of repository have been developed to support the self-archiving of research papers. Some of the earliest successful repositories were e-print services

developed to support communication in certain subject disciplines. For historical reasons, these tended to develop within disciplines with an existing tradition of sharing pre-prints or working papers, e.g. high-energy physics, computer science, and economics. Some of these services are now extremely well established in their disciplinary culture, e.g. the arXiv service (http://arxiv.org/) used by physicists. To supplement these discipline-based repositories, there has been a recent increased emphasis on the importance of institution-based repositories that collect, manage and make available digital assets on behalf of the institutions that produce them (Crow, 2002). In principle, these 'institutional repositories' have a much wider remit than research papers or e-prints, being seen as a way for institutions to emphasise their organisational commitment to the stewardship of all types of digital materials, including research and teaching materials produced by staff and students, institutional records and scientific data (Lynch, 2003). However, in practice, most institutional repositories have been deployed to support the 'self-archiving' of research publications or theses. Once such institution-based services have been developed and populated with content, tools like the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) can be used to link them into a global system of distributed interoperable repositories (Crow, 2002, p. 6). The JISC's Focus on Access to Institutional Resources (FAIR) programme was funded to explore the different ways that institutional content (e.g., research outputs, theses, images, museum objects) may be disclosed and shared using tools like the OAI-PMH.

There are three main ways in which institutional repositories might be able to support the UK model of research assessment. The simplest of these - and probably the easiest to realise in the short term - would be to use an institutional repository as a means of supporting the existing RAE processes, i.e. providing accurate metadata and the full-text of documents to the submission process and to unit of assessment panels. A second approach might take this further and use citation linking as the basis for a quantitative evaluation using bibliometric methods. For example, Harnad (2003) has proposed that every researcher should have a standardised electronic curriculum vitae with papers linked to a full-text version available in an institutional repository that could be used for citation linking and analysis. Thirdly, there may be additional Web-specific metrics that can be used to support evaluation, e.g. page downloads or Web link data. The remainder of this section will investigate these approaches in more detail.

4.1 Repositories in support of the RAE submission process

The process for the 2001 RAE required each unit being assessed to prepare a standardised submission for review by one of 68 discipline-based units of assessment. The submission needed to include information on numbers of staff and research students, grant income, research strategy, infrastructure and evidence of esteem. In the 2001 exercise, each member of staff included in the evaluation process could submit up to four research outputs (chiefly publications) for review. In order to facilitate the last part of this process, several UK universities devised systems - typically collections or databases of research publications - for this particular use. For example, in 2001 Leeds University Library provided information on publications, checked citation information and maintained a hard copy of all items submitted to the

RAE. They similarly propose that for the next exercise, the library - in collaboration with the university's Research Support Unit and Information Systems Services - will maintain a service called the University of Leeds Publications Database (ULPD). This will be a central repository for "information about all publications authored by University members of staff and postgraduates ... to store, locate and manage information about publications, generate publication reports, dynamically link publications to personal and departmental web pages and export publications data to bibliographic software packages" (http://www.leeds.ac.uk/library/teams/rae/).

It is clear that at least some of these functions could be undertaken or supported by an institutional repository. At the very least, the repository could be used to export authoritative metadata in standard formats to the RAE submission process, also to help manage digital copies of publications that could be made available to unit of assessment panels, on request.

4.2 Repositories and bibliometric analysis

A more sophisticated way of using institutional repositories to support the RAE would be to use citation-linking techniques as a means of automatically generating quantitative indicators that could be used for research evaluation. There have been a number of initiatives and projects that have investigated the use of citation linking in this context.

4.2.1 Automatic reference linking and citation analysis

One of the first proposals for a Web-based reference linking service for research papers was the 'universal citation database' proposed by Cameron (1997). He proposed a distributed database maintained by the organisations that produce scholarly work - a vision not dissimilar to that of institutional repositories linked by the OAI-PMH. The proposed system depended on unique identifiers linked to authoritative bibliographic data and the resulting database could be used for supporting searching as well as research evaluation.

CiteSeer (sometimes also known as ResearchIndex) is a digital library developed by the NEC Research Institute that uses a technique known as 'autonomous citation linking' to create a citation index from the full-text of downloaded papers in computer science (Lawrence, Giles & Bollacker, 1999). In order to do this, the system automatically downloads relevant documents, identifies common features like titles, authors and reference lists, then parses the citations to identify and extract information on authors, titles, year of publication, etc. In addition, the reference tags within the text of the paper are identified so that the search results interface can display the exact context of the citation in each paper. The unambiguous identification of citations, e.g. given differing or incomplete reference styles, can be problematic, so CiteSeer tested several algorithms to help improve this process. The resulting database of full-text papers and link information is then used to support the CiteSeer search interface, enabling the ranking of search results by the number of inward links (citations) and the navigation of both inward and outward links. The CiteSeer service available on the Web demonstrates these features on a growing corpus in the computer science domain (http://citeseer.ist.psu.edu/oai.html), currently (late 2004) searching over 700,000 documents. CiteSeer itself generates some of its

own bibliometric statistics, e.g. a list of the most cited authors and papers. The database has also been used for the predictive ranking of scientists by third parties (e.g., Feitelson & Yovel, 2004).

Citebase is a similar type of system, first developed as part of the Open Citation Project (http://opcit.eprints.org/), a joint project of the University of Southampton, Cornell University, and the arXiv repository (Hitchcock, et al., 2002). While CiteSeer builds its service upon papers made available through personal or research group Web pages, Citebase takes advantage of the slightly more structured infrastructure provided by e-print repositories. Citebase was specifically concerned with demonstrating the value of adding reference linking and citation services to e-print repositories, e.g. for helping to measure the research impact of individual papers or authors. Citebase works by harvesting both metadata records and bibliographic references from the full-text of e-prints stored in OAI-compliant repositories, initially arXiv, CogPrints, and BioMed Central. The association between these the records and the references then formed the basis of a citation database that could be used to count inward and outward links. In addition, Citebase collected data on the number of downloads (or hits) from the UK mirror of arXiv, which allowed the tentative comparison of download and citation frequency. The main Citebase search interface enables results to be ranked by date, by the number of citations received, or the number of downloads, thus allowing users to gain some indication of the research impact of papers while searching for e-prints. A full Citebase record includes bibliographic metadata, an abstract, the full reference list, the top five inward citations and co-citations, and a graph showing the known citation and download history [add figure?]. Citebase thus demonstrates some of the added value services that can be developed if e-prints are made freely available in OAI-compliant repositories.

Hitchcock, et al. (2002) freely admit that, when compared with mature citation services like Thomson ISI's Web of Science, Citebase and CiteSeer are both "in their infancy, covering diverse collections, having to work with inconsistent data formats, and trying to identify user preferences to optimise their features." However, both demonstrate the potential of developing innovative secondary services that provide automatic reference linking and citation analysis. Their existence has also inspired a provocative proposal for the replacement of the UK research assessment exercises by a continuous online assessment based on citation analysis.

4.2.2 Institutional repositories, citation analysis and research assessment

Harnad (2001) has proposed that the UK research assessment exercises could support the development of open access to research literature, "by mandating that all UK universities self-archive all their annual refereed research in their own eprint archives." Citation analysis tools could then be used to enable continuous online assessment, which could eventually replace the resource-intensive peer-review based evaluation process currently used by the RAEs.

In a more detailed elaboration of this proposal, Harnad, et al. (2003) first point to the statistical correlation between RAE rankings and citation counts found in bibliometric studies undertaken by Smith and Eysenck (2002) and Oppenheim and his colleagues (e.g., Oppenheim, 1995; 1996; Holmes & Oppenheim, 1998). From this they conclude that citation counts and the RAE are measuring broadly the same

thing. In order to take advantage of this, they propose that the higher education funding bodies should mandate all research-active staff to maintain continually updated online *curriculum vitae* known as a RAE-CV. These would include details of all refereed research papers, the full-text of which would be 'self-archived' in the university's institutional archive. This data could be retrieved and analysed by impact assessment services like Citebase to measure research productivity and impact. Criteria might include, for example, author and paper citation counts, journal impact factors, and the number of downloads or hits.

The proposal is forward-looking and provocative and, if implemented, would do much to promote the development of institutional repositories in the UK. Whether it would be a positive move forward for research assessment is more doubtful. A case in point would be books. The proposal suggests that authors could deposit the bibliographic references only so these could be used for citation linking and analysis, which seems to underestimate the difficulty (and time involved) in doing this. Also, the existence of different citation styles, e.g. of archival sources in historical works, will complicate the work of the citation linking tools. In any case, responses to the Roberts review of the RAE revealed little enthusiasm for a research evaluation exercise based purely on quantitative measures like citations. Despite the statistical correlation between citation counts and RAE scores found by Oppenheim, many scientists and scholars are extremely wary of solely using bibliometric analysis for the evaluation of research groups or individuals.

4.3 Usage statistics

Citations apart, another kind of metric that could be used to support the evaluation of research would be usage statistics. Web log data is routinely collected and analysed in e-commerce contexts, e.g. to shed light on the utilisation of Web sites, for analysing navigation patterns and other aspects of online behaviour (Berendt & Spiliopoulou, 2000). The evidence gathered in this way can be used for improving Web site navigation features or for personalisation. Web servers can gather user log data automatically, but the limitations of the data collected means that it needs very careful interpretation (Nicholas, et al., 1999). Various data mining methods have been developed to overcome these limitations (Spiliopoulou, 2000; Spiliopoulou & Pohle, 2001). Web usage data is most often used to support e-commerce applications, but is also used by research libraries, e.g. to evaluate the effectiveness of information services and of user behaviour (e.g., Bracke, 2004). Libraries and library consortia also collect e-journal usage statistics, e.g. data on searches undertaken, abstracts viewed, full-texts displayed or downloaded, although there is still a need for standardised information to be provided by publishers (Luther, 2001). Some standards on the production of statistics exist, e.g. the guidelines produced by the International Coalition of Library Consortia (ICOLC), the COUNTER (Counting Online Usage of Networked Electronic Resources) project code of practice (http://www.projectcounter.org/code_practice.html), and the revision of ANSI/NISO Z39.7 - 1995 currently being undertaken by The US National Information Standards Organization (NISO). However, the data collected still needs very careful analysis.

Some pioneering studies have demonstrated the potential use of usage data as a quantitative measure for the evaluation of research. For example, Kurtz, et al. (2002;

2004) compared article usage (reads) from the NASA Astrophysics Data System with citation data. Statistical analysis found that while reads and cites have different properties, both are measures of the usefulness of a paper. Usage and citation data can be combined to measure the productivity of individual researchers, although Kurtz and his colleagues warn that this is not "accurate enough to preclude the necessity of careful and direct examination of the record of the person being evaluated." They conclude that usage information is a more direct measure of current usefulness and is well suited for bibliometric use. Perneger (2004) also found a positive correlation between 'hit counts' and future citation counts in a study of the *BMJ*, concluding that the publication of download statistics by online journals should be encouraged.

E-print repositories are well placed to capture some usage information. Like Kurtz, and his colleagues, Harnad (2003) has argued that usage (download) data should be seen as "usage (and hence impact) measures in their own right." Brody and Harnad (2004) used download data from the arXiv service to investigate the relationship between usage and citation rates. Papers were harvested from arXiv and Citebase used to generate citation links and an impact score for each. Web log data was also collected from the UK mirror of arXiv, filtered and analysed to see whether there was any correlation with the citation rate. The results were very promising, suggesting that short-term 'download impact' could predict - to some extent - medium-term citation rates.

The utilisation of usage data in bibliometric research is at a very early stage. Current research suggests that while there are significant differences between usage and citation data, they could be used in combination to support research evaluation.

4.4 Webometrics

The study of the link structure of the World Wide Web has become the subject of much recent research in statistical physics, computer science, and information science (e.g., Barabási, 2001; Kleinberg & Lawrence, 2001). One practical application of this research has been the use of hyperlink data to support Web information retrieval and result ranking or clustering (Henzinger, 2001). The most well known example of this is the search service Google (http://www.google.com/), whose PageRank algorithm ranks Web pages according to the number and nature of incoming links (Brin & Page, 1998). Similar ideas were implemented in an experimental search engine developed by IBM Research, which used a 'hyperlinked-induced topic search' algorithm to help identify pages with good quality content ('authorities') and those Web pages that link to many such authorities ('hubs'), the latter of which being assumed to provide good starting points for browsing (Kleinberg, 1999; Chakrabarti, et al., 1999).

These services use bibliometric-type techniques to support retrieval and the ranking of results. Information scientists have also used the same techniques to explore the nature of online communication - a type of research sometimes known as 'Webometrics' (Almind & Ingwersen (1997; Björneborn & Ingwersen, 2001). Ingwersen (1998), for example, introduced the concept of Web Impact Factors (WIFs), broadly based on the total of external or internal pages that point to a given

site or domain, divided by the total number of pages in that site or domain, at a particular moment in time.

Much Webometric research has focused on university Web sites, e.g. generating different kinds of WIF and comparing them with existing measures of research output or quality. The results of these studies have been varied. For example, Smith (1999) looked at Australasian universities, finding no significant correlation with research output as measured by publication counts. On the other hand, Thelwall (2001) was able to show *some* association between external Web impact factors for 25 UK university web sites and institution-level research ranking data derived from the 1996 RAE. Refinements in methodology, largely related to the development of new document models, meant that follow-up research on 108 UK Universities was able to reveal a highly statistically significant correlation between Web inlinks and research productivity, as measured by RAE 2001 rankings (Thelwall, 2002; Thelwall & Wilkinson, 2003). Webometric analysis at the departmental level has also had mixed results. Thomas and Willett (2000) compared WIFs with citation data for UK library and information science departments, but found no significant correlation with the rankings from RAE 1996. More recently, however, a study of UK computer science departments showed that the number of Web inlinks correlated with research productivities when the size of departments was taken into account (Li, *et al*., 2003).

These results indicate that there is some association between inward Web linking behaviour and measures of research productivity. The exact nature of this association is a matter for future research, although Thelwall (2004, p. 127) has argued that any "correlations found should be treated with caution and do not prove a cause and effect relationship." Thelwall (2001, p. 1166) first speculated that some Web impact measures might be reflecting the reputation of a university and its staff, rather than the quality of the content. However, a more recent study, comparing Web link data with RAE 2001 rankings, suggests that high numbers of inlinks may just be a reflection of the *quantity* of Web content available. Thelwall and Harries (2004) found no evidence that higher rated scholars produced Web content with a significantly higher impact, only that they produced more. This finding is important, meaning that Web link data should not be used for the evaluation of research groups or individuals, and with extreme care at higher levels of granularity.

Current Webometric studies into university Web linking behaviour are interesting because of the potential insight that they give into scholarly linking behaviour. Studies suggest that the Web is used mainly for informal communication, e.g. a study of random links between UK university Web sites found that while the majority (around 90%) were related to scholarly activities (research and teaching), very few (0.5%) were traditional references to research papers (Wilkinson, et al., 2003). There is a need for more research into understanding linking behaviour. A lot of effort has been made to develop and improve the methodologies used in Webometric research, although it is still a relatively young discipline and the Web is a particularly challenging and fluid environment to study (e.g. Egghe, 2000). Institutional repositories would provide an additional source of data for Webometric research, although the linking behaviour would have more in common with traditional bibliometric analysis of scholarly publications. It might be possible to develop impact measures for individual repositories, although it would have to be made clear that these should not be regarded as a surrogate for the research quality of institutions.

5. The role of ePrints UK

In terms of the OAI-PMH, ePrints UK is developing a demonstration 'service provider' that will harvest and aggregate metadata from multiple UK repositories. In addition, experimental Web services will be developed to help provide authoritative metadata on personal or organisational names, some subject classification, and to parse bibliographic references to support citation linking.

The potential role that any operational ePrints UK system could have with regard to supporting the UK research assessment exercises would be limited by its role as an OAI service provider. So, for example, an ePrints UK would not by itself be able to directly support the RAE activities of individual institutions, e.g. in generating metadata in suitable formats for submission to the exercise. Provided with a means of identifying which items in UK repositories were submitted to the RAE, a system like ePrints UK *could* aggregate such resources to facilitate their use by panels, but it remains to be seen whether this would be useful in practice.

EPrints UK would be better placed to support the collection of citation data. The project has explored the automatic generation of citation linking data through Web Services, and this could form the basis of gathering citation data on the model provided by Citebase which could possibly contribute bibliometric information for use by those evaluating research. The potential for this needs to be further explored

The ePrints UK aggregator could also collect other data, e.g. information on article downloads or the type of searches made. This kind of data is problematic to analyse, but it arguably should be somehow used as a bibliometric measure in its own right, probably in combination with other metrics. It is perhaps worth noting that the ePrints UK aggregator does not currently harvest usage data from data providers. If this is seen as a desirable property, thought will have to be given to how this should be done, e.g. possibly through the OAI-PMH.

6. Conclusions

The research assessment exercises remain a subject of controversy within UK higher education institutions. While they appear, in Adams (2002) words, to have become a "pervasive driver of excellence," the increased level of 'research concentration' following the 2001 exercise has prompted some to suggest that they have now fulfilled their original objective and that the 2008 RAE should be the final one (MacLeod, 2004).

The Roberts review considered the desirability of giving performance metrics an increased role in the exercises but in consultation it found little support in the community for a RAE largely based on such measures. RAE 2008 will, therefore, still be based on peer-review, although panels are encouraged to use metrics to support decisions where appropriate.

Given the existence of enough content, institutional repositories would be able to automatically generate some additional bibliometric data, e.g. measures of citations or usage, which could be used by RAE panels to support peer-review-based evaluation. However the applicability of this data across disciplines still needs to be tested. The proposal for RAE CVs linked to institutional repositories that could then

be used for bibliometric analysis and continuous research evaluation (Harnad, et al., 2003) is innovative but would be difficult to implement across all disciplines. It may be a useful way of persuading institutions to develop repositories and academics to deposit research outputs in them, but any continuous research assessment based *purely* on bibliometric analysis is likely to be unacceptable to many academics and probably have other unintended consequences. Despite the known disadvantages of peer review, academics remain suspicious of performance metrics, believing them "too crude a means to assess the quality of research" (Roberts, 2003, p. 82). Despite the growing sophistication of bibliometric studies (Borgman & Furner, 2002), many scientists and scholars would agree with Brenner (1995) when he argued that before "we develop a pseudoscience of citation analysis, we should remind ourselves that what matters absolutely is the scientific content of a paper and that nothing will substitute for either knowing it or reading it."

The one area where institutional repositories could already support research assessment is in helping to provide a means to generate, store and manage information about research outputs that can be used to support the submission and review process. This is, however, the area in which service providers like ePrints UK have the least to offer. Apart from this key area, using the RAE as a means of promoting the development of institutional repositories may be counter-productive. When talking about the institutional control of resources, Lynch (2003) has argued that institutional repositories will succeed only if they are responsive to the needs of campus communities and scholarship, warning that where institutions "try to *enforce* behavioral or cultural changes -- and particularly controversial ones -- within the campus community they will and should fail."

Acknowledgements

UKOLN is funded by the Museums, Libraries and Archives Council (MLA), the Joint Information Systems Committee (JISC) of the UK higher and further education funding councils, as well as by project funding from JISC, the European Union and other organisations. UKOLN also receives support from the University of Bath, where it is based. This study has been produced for ePrints UK, a project funded by the JISC as part of its Focus on Access to Institutional Resources (FAIR) Programme.

References

Adam, D. (2002). "The counting house." *Nature*, 415, 726-729.

Adams, J. (1998). "Benchmarking international research." *Nature*, 396, 615-618.

Adams, J. (2002). "Research assessment in the UK." *Science*, 296, 805.

Adams, J., Bailey, T., Jackson, L., Scott, P., Pendlebury, D., & Small, H. (1997). "Benchmarking of the international standing of research in England." Leeds: University of Leeds, Centre for Policy Studies in Education. Retrieved December 2, 2004, from http://www.evidenceuk.com/downloads/benchmark-report.pdf

Adams, J., Cook, N., Law, G., Marshall, S., Mount, D., Smith, D., Wilkinson, D., Bayers, N., Pendlebury, D., Small, H., & Stephenson, J. (2000). "The role of selectivity and the characteristics of excellence." Leeds: University of Leeds, Higher Education Policy Unit. Retrieved December 2, 2004, from http://www.evidenceuk.com/downloads/selectivity-report.pdf

Adams, J., Jackson, L., Law, G., Mount, D., Reeve, N., Smith, D. N., & Wilkinson, D. (2002). "Maintaining research excellence and volume." Leeds: Evidence, Ltd. Retrieved December 2, 2004, from http://www.evidenceuk.com/downloads/maint-research-exellence.pdf

Ball, D. F. (1997). "Quality measurement as a basis for resource allocation: research assessment exercises in United Kingdom universities." *R&D Management*, 27(3), 281-289.

Ball, D. F., & Butler, J. (2004). "The implicit use of business concepts in the UK Research Assessment Exercise." *R&D Management*, 34(1), 87-97.

Barabási, A.-L. (2001). "The physics of the Web." *Physics World*, 14(7), 33-38. Retrieved December 2, 2004, from http://physicsWeb.org/article/world/14/7/9

Bar-Ilan, J. (2001). "Data collection methods on the Web for informetric purposes: a review and analysis." *Scientometrics*, 50(1), 7-32.

Bence, V., & Oppenheim, C. (2001). "Journals, scholarly communication and the RAE: a case study of the business and management sector." *Serials*, 14(3), 265-272.

Berendt, B. & Spiliopoulou, M. (2000). "Analysis of navigation behaviour in Web sites: integrating multiple information systems." *VLDB Journal*, 9, 56-75.

Bessant, J., Birley, C., Cooper, C., Dawson, S., Gennard, S., Gardiner, M., Gray, A., Jones, P., Mayer, C., McGee, J., Pidd, M., Rowley, G., Saunders, J., & Stark, A. (2003). "The state of the field in UK management research: reflections of the Research Assessment Exercise (RAE) panel." *British Journal of Management*, 14, 51-68.

Borgman, C. L., & Furner, J. "Scholarly communication and bibliometrics." *Annual Review of Information Science and Technology*, 36, 3-72.

Bracke, P. J. (2004). "Web usage mining at an academic health sciences library: an exploratory study" *Journal of the Medical Library Association*, 92(4), 421-428.

Brenner, S. (1995). "Loose ends." *Current Biology*, 5(5), 568.

Brody, T, & Harnad, S. (2004). "Earlier Web usage statistics as predictors of later citation impact." Retrieved December 2, 2004, from http://www.ecs.soton.ac.uk/~harnad/Temp/timcorr.doc [is this a public document?]

Brin, S., & Page, L. (1998). "The anatomy of a large-scale hypertextual Web search engine." *Computer Networks and ISDN Systems*, 30(1-7), 107-117.

Cameron, R. D. (1997). "A universal citation database as a catalyst for reform in scholarly communication." *First Monday*, 2(4). Retrieved December 2, 2004, from http://www.firstmonday.dk/issues/issue2_4/cameron/

Chakrabarti, S., Dom, B.E., Kumar, S.R., Raghavan, P., Rajagopalan, S., Tomkins, A., Gibson, D. & Kleinberg, J. (1999). "Mining the Web's link structure." *Computer*, 32(8), 60-67.

Clark, D. (2002, 6 December). "Precocious teenager will benefit from scholarship." *Times Higher Education Supplement*, 1567, 17.

Crow, R. (2002). *The case for institutional repositories: a SPARC position paper*. Washington, D.C.: Scholarly Publishing & Academic Resources Coalition. Retrieved December 2, 2004, from http://www.arl.org/sparc/IR/ir.html

Department for Education and Skills. (2003a). "The future of higher education." Cm. 5735. London: The Stationery Office, January. Retrieved December 2, 2004, from http://www.dfes.gov.uk/hegateway/strategy/hestrategy/

Department for Education and Skills. (2003b). "The future of higher education: commentary on responses received on the Higher Education White Paper and paper on widening participation." London: DfES, July. Retrieved December 2, 2004, from http://www.dfes.gov.uk/hegateway/strategy/hestrategy/

Deutsche Forchungsgemeinschaft. (1998). *Recommendations of the Commission on Professional Self Regulation in Science - Proposals for safeguarding good scientific practice*. Retrieved December 2, 2004, from http://www.dfg.de/aktuelles_presse/reden_stellungnahmen/ download/self_regulation_98.pdf

Doyle, J. R., & Arthurs, A. J. (1995). "Judging the quality of research in business schools: the UK as a case study." *Omega*, 23(3), 257-270.

Doyle, J. R., Arthurs, A. J., Green, R. H., McAulay, L., Pitt, M. R., Bottomley, P. A., & Evans, W. (1996). "The judge, the model of the judge, and the model of the judge as judge: analyses of the UK 1992 research assessment exercise data for business and management studies." *Omega*, 24(1), 13-28.

Egghe, L. (2000). "New informetric aspects of the Internet: some reflections - many problems." *Journal of Information Science*, 26(5), 329-335.

Elkin, J. (2002). "The UK Research Assessment Exercise 2001." 68th IFLA Council and General Conference, Glasgow, UK, 18-24 August. Retrieved December 2, 2004, from http://www.ifla.org/IV/ifla68/papers/046-156e.pdf

Elton, L. (2003). "Goodhart's Law and performance indicators in higher education." Retrieved December 2, 2004 from the Hong Kong Institute of Education Web site: http://www.ied.edu.hk/loap/Goodhart%27%20law.pdf

Evidence Ltd. (2003a). "Funding research diversity: the impact of further concentration of university research performance and regional research capacity." Retrieved December 2, 2004, from http://www.universitiesuk.ac.uk/diversity/

Evidence Ltd. (2003b). "PSA target metrics for the UK research base." Retrieved December 2, 2004, from http://www.ost.gov.uk/policy/psa_target_metrics.pdf

Farrant, J., Billing, D., & Temple, P. (2003). "Operational review of the Research Assessment Exercise 2001." Milton Keynes: Universitas. Retrieved December 2, 2004, from http://www.ra-review.ac.uk/reports/

Fazackerley, A. (2004, 21 May). "Staff at risk in RAE run-up." *Times Higher Education Supplement*, 1641, 1.

Fazackerley, A. (2004, 26 November). "Rescue for top-rated staff only." *Times Higher Education Supplement*, 1669, 1.

Gambrill, E. (2002, 6 December). "More clarity and less secrecy will counter bias." *Times Higher Education Supplement*, 1567, 16.

Garfield, E. (1983). *Citation indexing: its theory and application in science, technology, and humanities*. Philadelphia, Penn.: ISI Press.

Garfield, E. (1996). "How can impact factors be improved?" *BMJ*, 313, 411-413.

Garfield, E. (1998). "From citation indexes to informetrics: is the tail now wagging the dog? *Libri*, 48, 67-80.

Garfield, E. (1999). "Journal impact factor: a brief review." *Canadian Medical Association Journal*, 161(8), 979-980.

Garner, M. (2002, 6 December). "Tired, hard-up RAE seeks saviour with light touch." *Times Higher Education Supplement*, 1567, 16-17.

Georghiou, L., Howells, J., Rigby, J., Glynn, S., Butler, J., Cameron, H., Cunningham, P., Ausadamongkol, K., Thomas, D., Salazar, A., Barker, K., & Reeve, N. (2000). "Impact of the Research Assessment Exercise and the future of quality assurance in the light of changes in

the research landscape." Manchester: University of Manchester, PREST. Retrieved December 2, 2004, from http://www.hefce.ac.uk/Research/review/consult/raec.pdf

Geuna, A., & Martin, B. R. (2003). "University research evaluation and funding: an international comparison." *Minerva*, 41, 277-304.

Glänzel, W., & Moed, H. F. (2002). "Journal impact measures in bibliometric research." *Scientometrics*, 53(2), 171-193.

Glover, E. J., Lawrence, S., Gordon, M. D., Birmingham, W. P., & Giles, S.L. (2001). "Web search - your way." *Communications of the ACM*, 44(12), 97-102.

Goddard, A. (2003, 30 May). "'Radical change' planned for RAE." *Times Higher Education Supplement*, 1591, 7.

Goddard, A., & Davis, C. (2003, 26 September). "Clear improvement?" *Times Higher Education Supplement*, 1608, 12.

Harnad, S. (2001). "The self-archiving initiative." *Nature*, 410, 1024-1025. Retrieved December 2, 2004, from http://eprints.ecs.soton.ac.uk/archive/00005947/

Harnad, S. (2001, 18 May). "Why I ... think research access, impact and assessment are linked." *Times Higher Education Supplement*, 1487, 16. Retrieved December 2, 2004, from http://eprints.ecs.soton.ac.uk/archive/00005950/

Harnad, S. (2003, 6 June). "Why I ... believe that all UK research output should be online." *Times Higher Education Supplement*, 1592, 16.

Harnad, S., Carr, L., Brody, T., & Oppenheim, C. (2003). "Mandated online RAE CVs linked to university eprint archives: enhancing UK research impact and assessment." *Ariadne*, 35. Retrieved December 2, 2004, from http://www.ariadne.ac.uk/issue35/harnad/

Higher Education Funding Council for England. (2003). "Funding higher education in England: how HEFCE allocates its funds." HEFCE 2003/29. Bristol: HEFCE. Retrieved December 2, 2004, from http://www.hefce.ac.uk/Pubs/hefce/2003/03_29.htm

Higher Education Funding Council for England. (2004). "Recurrent grants for 2004-05." HEFCE 2004/12. Bristol: HEFCE. Retrieved December 2, 2004, from http://www.hefce.ac.uk/Pubs/hefce/2004/04_12/

Higher Education Funding Council for England, Scottish Higher Education Funding Council, Higher Education Funding Council for Wales, & Department for Employment and Learning Northern Ireland. (2004). "RAE2008: initial decisions by the UK funding bodies." RAE 01/2004. Retrieved December 2, 2004, from http://www.rae.ac.uk/pubs/2004/01/

Hinze, S. (2001). "Bibliometric analysis of Norwegian research activities." Reports in the evaluation of the Research Council of Norway, background report, no. 2. Retrieved December 2, 2004, from http://www.technopolis.co.uk/reports/RCN/RCN_BR02.pdf

Hitchcock, S., & Hall, W. (2001). "How dynamic e-journals can interconnect open access archives." In: Hubler, A., Linde, P. & Smith, J.W.T. (eds.) *Electronic publishing '01: 2001 in the digital publishing odyssey*. Amsterdam: IOS Press, 183-193. Retrieved December 2, 2004, from http://eprints.ecs.soton.ac.uk/archive/00005982/

Hitchcock, S., Carr, L., Jiao, Z., Bergmark, D., Hall, W., Lagoze, C., & Harnad, S. (2000). "Developing services for open eprint archives: globalisation, integration and the impact of links." In: *Proceedings of the 5th ACM Conference on Digital Libraries*. New York: ACM Press, 143-151.

Hitchcock, S., Bergmark, D., Brody, T., Gutteridge, C., Carr, L., Hall, W., Lagoze, C., & Harnad, S. (2002). "Open citation linking: the way forward." *D-Lib Magazine*, 8(10). http://www.dlib.org/dlib/october02/hitchcock/10hitchcock.html

Holmes, A., & Oppenheim, C. (2001). "Use of citation analysis to predict the outcome of the 2001 Research Assessment Exercise for Unit of Assessment (UoA) 61: Library and

Information Management." *Information Research*, 6(2), January. Retrieved December 2, 2004, from http://informationr.net/ir/6-2/paper103.html

Horrobin, D. F. (1996). "Peer review of grant applications: a harbinger for mediocrity in clinical research?" *The Lancet*, 348, 1293-1295.

House of Commons, Select Committee on Science and Technology. (2002). "The Research Assessment Exercise." HC 507. London: The Stationery Office. Retrieved December 2, 2004, from http://www.publications.parliament.uk/pa/cm200102/cmselect/cmsctech/507/50702.htm

House of Commons, Select Committee on Science and Technology. (2004a). "Scientific publications: free for all?" HC 399-I. London: The Stationery Office. Retrieved December 2, 2004 from http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/cmsctech.htm

House of Commons, Select Committee on Science and Technology. (2004b). "Research Assessment Exercise: a re-assessment." HC 586. London: The Stationery Office. Retrieved December 2, 2004 from http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/cmsctech. htm

House of Commons, Select Committee on Science and Technology. (2004c). "Responses to the Committee's tenth report, Session 2003-04, scientific publications, free for all?" HC 1200. London: The Stationery Office. Retrieved December 2, 2004 from http://www.parliament.the-stationery-office.co.uk/pa/cm200304/cmselect/cmsctech/cmsctech.htm

House of Commons, Select Committee on Science and Technology. (2004d). "Research Assessment Exercise: a re-assessment: Government response to the Committee's eleventh report of Session 2003-04." HC 34. London: The Stationery Office. Retrieved December 2, 2004 from http://www.parliament.the-stationery-office.co.uk/pa/cm/cmsctech.htm

Jefferson, T., Alderson, P., Wager, E., & Davidoff, F. (2002). "Effects of editorial peer review: a systematic review." *JAMA*, 287, 2784-2786.

Jefferson, T., Wager, E., & Davidoff, F. (2002). "Measuring the quality of editorial peer review." *JAMA*, 287, 2786-2790.

Jennings, C. (1998). "Citation data: the wrong impact?" *Nature Neuroscience*, 1(8), 641-642.

Johnes, J. (1996). "Performance assessment in higher education in Britain." *European Journal of Operational Research*, 89(1), 18-33.

Johnes, J., & Taylor, J. (1992). "The 1989 research selectivity exercise: a statistical analysis of differences in research rating between universities at the cost centre level." *Higher Education Quarterly*, 46(1), 67-87.

Johnes, J., Taylor, J., & Francis, B. (1993). "The research performance of UK universities: a statistical analysis of the results of the 1989 research selectivity exercise." *Journal of the Royal Statistical Society A*, 156(2), 271-286.

Kaltenborn, K.-F., & Kuhn, K. (2003). "Der Impact-Faktor als Parameter zur Evaluation von Forcherinnen/Forschern und Forschung." *Medizinische Klinik*, 98(3), 153-169.

King, D.A. (2004). "The scientific impact of nations." *Nature*, 430, 311-316.

King, J. (1987). "A review of bibliometric and other science indicators and their role in research evaluation." *Journal of Information Science*, 13, 261-276.

Kleinberg, J. (1999). "Authoritative sources in a hyperlinked environment." *Journal of the ACM*, 46(5), 604-632.

Kleinberg, J., & Lawrence, S. (2001). "The structure of the Web." *Science*, 294, 1849-1850.

Knight, P. (2002, 6 December). "Gladiatorial contest is the victim of its own success." *Times Higher Education Supplement*, 1567, 16.

Kokko, H. (1999). "What do impact factors tell us?" *Trends in Ecology and Evolution*, 14(10), 382-384.

Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C. S., Demleitner, M., Murray, S. S., Martimbeau, N., & Elwell, B. (2004). "The NASA Astrophysics Data System: sociology, bibliometrics, and impact." Retrieved December 2, 2004, from http://cfa-www.harvard.edu/~kurtz/jasis-abstract.html

Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C. S., Thompson, D. M, Bohlen, E. H., & Murray, S. S. (2002). "The NASA Astrophysics Data System: obsolescence of reads and cites." In: Corbin, B. G., Bryson, E. P., & Wolf, M., eds., *Library and Information Services in Astronomy IV (LISA IV): proceedings of a conference held at Charles University, Prague, Czech Republic, July 2-5, 2002*. Retrieved December 2, 2004, from http://www.eso.org/gen-fac/libraries/lisa4/Kurtz.pdf

Lawrence, S. (2001). "Free online availability substantially increases a paper's impact." *Nature*, 411, 521. Retrieved December 2, 2004, from http://www.nature.com/nature/debates/e-access/Articles/lawrence.html

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

Li, X., Thelwall, M., Musgrove, P., & Wilkinson, D. (2003). "The relationship between the WIFs or inlinks of computer science departments in UK and their RAE ratings or research productivities in 2001." *Scientometrics*, 57(2), 239-255.

Luther, J. (2001). *White paper on electronic journal usage statistics*, 2nd ed. Washington, D.C.: Council on Library and Information Resources. Retrieved December 2, 2004, from http://www.clir.org/pubs/abstract/pub94abst.html

Lynch, C. A. (2003). "Institutional repositories: essential infrastructure for scholarship in the digital age." *ARL Bimonthly Report*, 226, February. Retrieved December 2, 2004, from http://www.arl.org/newsltr/226/ir.html

May, R. M. (1997). "The scientific wealth of nations." *Science*, 275, 793-796.

May, R. M. (1998). "The scientific investments of nations." *Science*, 281, 49-51.

McKay, S. (2003). "Quantifying quality: can quantitative data ('metrics') explain the 2001 RAE ratings for social policy and administration?" *Social Policy & Administration*, 37(5), 444-467.

MacLeod, D. (2004, 9 November). "This could be the last time." *The Guardian*, [pp.]. Retrieved December 2, 2004, from http://education.guardian.co.uk/egweekly/story/0,5500,1346153,00.html

Merton, R. H. (1968) "The Matthew Effect in science." *Science*, 159, 56-63.

Moed, H. F. (2002). "The impact-factors debate: the ISI's uses and limits." *Nature*, 415, 731-732.

Nature. (1997a). "Reviews of UK university research 'have helped to raise standards'." *Nature*, 385, 3.

Nature. (1997b). "Research assessment's ups and downs." *Nature*, 385, 99.

Nature. (2002). "Errors in citation statistics." *Nature*, 415, 101.

Nicholas, D., Huntington, P., Lievesley, N., & Withey, R. (1999). "Cracking the code: Web log analysis." *Online Information Review*, 23(5), 263-269.

Norris, M., & Oppenheim, C. (2003). "Citation counts and the Research Assessment Exercise V: archaeology and the 2001 RAE." *Journal of Documentation*, 59(6), 709-730.

Oehm, V., & Lindner, U. K. (2002, 31 May). "Umstritten, aber etabliert - der Impact Factor." *Deutsches Ärzteblatt*, 99(22), A1489-A1490. Retrieved December 2, 2004, from http://www.aerzteblatt.de/v4/archiv/artikel.asp?id=31829

Office of Science and Technology. (2003). "The sustainability of university research: a consultation on reforming parts of the dual support system." London: Department of Trade

and Industry. Retrieved December 2, 2004, from http://www.ost.gov.uk/policy/universityresearch.pdf

Oppenheim, C. (1995). "The correlation between citation counts and the 1992 Research Assessment Exercise ratings for British library and information science university departments." *Journal of Documentation*, 51(1), 18-27.

Oppenheim, C. (1996). "Do citations count? Citation indexing and the Research Assessment Exercise (RAE)." *Serials*, 9(2), 155-161.

Oppenheim, C. (1997). "The correlation between citation counts and the 1992 Research Assessment Exercise ratings for British research in genetics, anatomy and archaeology." *Journal of Documentation*, 53(5), 477-487.

Perneger, T. V. (2004). "Relation between online 'hit counts' and subsequent citations: prospective study of research papers in the *BMJ*." *BMJ*, 329, 546-547.

Punter, J., Morphet, J., Lloyd, M. G., & Boyle, R. (2002). "Research forum: reflections on the 2001 United Kingdom Research Assessment Exercise for town and country planning." *Town Planning Review*, 73(3), 333-360.

RAE Team. (1999). "Research Assessment Exercise 2001: assessment panels' criteria and working methods." RAE Circular 5/99. Bristol: RAE Team. Retrieved December 2, 2004, from http://www.hero.ac.uk/rae/Pubs/5_99/

Roberts, C. (1999). "Possible bias due to panel membership in the 1996 Research Assessment Exercise." University of Manchester, School of Epidemiology and Health Sciences. Retrieved December 2, 2004, from http://www.stopford.man.ac.uk/epidem/biostats/raepanel.rtf

Roberts, G. (2003). "Review of research assessment: report by Sir Gareth Roberts to the UK funding bodies." Joint Funding Bodies' Review of Research Assessment, consultation document, May. Retrieved December 2, 2004, from http://www.ra-review.ac.uk/reports/roberts.asp

Rothwell, N. (2004, 30 April). "'The peer-review system is creaking under the weight of work. How many grants can we assess and still find time to mark several hundred essays?'" *Times Higher Education Supplement*, 1638, 13.

Royal Society. (2003). "Supporting basic research in science and engineering: a call for a radical review of university research funding in the UK." Policy document 25/03. London: The Royal Society.

Seglen, P. O. (1992). "The skewness of science." *Journal of the American Society for Information Science*, 43(9), 628-638.

Seglen, P. O. (1997). "Why the impact factor of journals should not be used for evaluating research." *BMJ*, 314, 498-502.

Seglen, P. O. (1998). "Citation rates and journal impact factors are not suitable for evaluation of research." *Acta Orthopaedica Scandinavica*, 69(3), 224-229.

Smith, A. (1999). A tale of two Web spaces: comparing sites using Web impact factors." *Journal of Documentation*, 55(5), 577-592.

Smith, A., & Eysenck, M. (2002). "The correlation between RAE ratings and citation counts in psychology." Technical Report, Department of Psychology, Royal Holloway, University of London. Retrieved December 2, 2004, from http://cogprints.ecs.soton.ac.uk/archive/00002749/

Spiliopoulou, M. (2000). "Web usage mining for Web site evaluation." Communications of the ACM, 43(8), 127-134.

Spiliopoulou, M. & Pohle, C. (2001). "Data mining for measuring and improving the success of Web sites." *Data Mining and Knowledge Discovery*, 5, 85-114.

Strathern, M. (1997). "'Improving ratings:' audit in the British university system." *European Review*, 5, 305-321.

Strogatz, S. (2001). "Exploring complex networks." Nature, 410, 268-276.

Styles, D. R. (2002). "Higher education funding council (HEFC) methods in the 1990s: national and regional developments and policy implications." *Public Administration*, 80(4), 711-731.

Taylor, J. (1995). "A statistical analysis of the 1992 research assessment exercise." *Journal of the Royal Statistical Society A*, 158(2), 241-261.

Thelwall, M. (2001). "Extracting macroscopic information from Web links." *Journal of the American Society for Information Science and Technology*, 52(13), 1157-1168.

Thelwall, M. (2002). "Conceptualizing documentation on the Web: an evaluation of different heuristic-based models for counting links between university web sites." *Journal of the American Society for Information Science and Technology*, 53(12), 995-1005.

Thelwall, M. (2004). "Methods for reporting on the targets of links from national systems of university Web sites." *Information Processing and Management*, 40, 125-144.

Thelwall, M., & Harries, G. (2003). "The connection between the research of a university and counts of links to its Web pages: an investigation based upon a classification of the relationships of pages to the research of the host university." *Journal of the American Society for Information Science and Technology*, 54(7), 594-602.

Thelwall, M., & Harries, G. (2004). "Do better scholars' Web publications have significantly higher online impact?" *Journal of the American Society for Information Science and Technology*, 55(2), 149-159.

Thelwall, M., & Wilkinson, D. (2003). "Three target documentation range metrics for university Web sites." *Journal of the American Society for Information Science and Technology*, 54(6), 489-496.

Thomas, O., & Willett, P. (2000). "Webometric analysis of departments of librarianship and information science." *Journal of Information Science*, 26(6), 421-428.

Times Higher Education Supplement. (2003, 21 November). "Time running out for dual support." *Times Higher Education Supplement*, 1616, 14.

Times Higher Education Supplement. (2004, 21 May). "Better the RAE devil we know?" *Times Higher Education Supplement*, 1641, 18.

Tomlinson, S. (2000). "The research assessment exercise and medical research." *BMJ*, 320, 636-639.

Tunzelmann, F. von, & Kraemer Mbula, E. (2003). "Changes in research assessment practices in other countries since 1999: final report." Brighton: University of Sussex, SPRU Science and Technology Policy Research. Retrieved December 2, 2004, from http://www.ra-review.ac.uk/reports/

Tysome, T. (2004, 2 April). "Next RAE may demote 5*s." *Times Higher Education Supplement*, 1634, 7.

University Grants Committee. (1984). "A strategy for higher education into the 1990s." London: Her Majesty's Stationery Office.

Van de Sompel, H., & Beit-Arie, O. (2001). "Open linking in the scholarly information environment using the OpenURL Framework." *D-Lib Magazine*, 7(3). Retrieved December 2, 2004, from http://www.dlib.org/dlib/march01/vandesompel/03vandesompel.html

Van de Sompel, H., & Hochstenbach, P. (1999). "Reference linking in a hybrid library environment: part 1, frameworks for linking." *D-Lib Magazine*, 5(4). Retrieved December 2, 2004, from http://www.dlib.org/dlib/april99/van_de_sompel/04van_de_sompel-pt1.html

van Leeuwen, T. 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 research performance." *Scientometrics*, 51(1), 335-346.

van Raan, A. F. J. (2004). "Fatal attraction: ranking of universities by bibliometric methods." Retrieved December 2, 2004, from http://www.cwts.nl/cwts/RankingScientom.pdf

Warner, J. (1998). "The public reception of the Research Assessment Exercise 1996." *Information Research*, 3(4), April. Retrieved December 2, 2004, from http://informationr.net/ir/3-4/paper45.html

Warner, J. (2003). "Citation analysis and research assessment in the United Kingdom." *Bulletin of the American Society for Information Science and Technology*, 30(1), 26-27. Retrieved December 2, 2004, from http://www.asis.org/Bulletin/Oct-03/

Wellcome Trust. (2003). *Wellcome Trust position statement in support of open access publishing*. Retrieved December 2, 2004, from http://www.wellcome.ac.uk/doc% 5Fwtd002766.html

Wessely, S. (1998). "Peer review of grant applications: what do we know?" *The Lancet*, 352, 301-305.

Willmott, H. (2003). "Commercialising higher education in the UK: the state, industry and peer review." *Studies in Higher Education*, 28(2), 129-141. Retrieved December 2, 2004, from http://www.jims.cam.ac.uk/people/faculty/pdfs/willmott_commercialising_higher.pdf

Wilkinson, D., Harries, G., Thelwall, M., & Price, L. (2003). "Motivations for academic Web site interlinking: evidence for the Web as a novel source of information on informal scholarly communication." *Journal of Information Science*, 29(1), 49-56.

Ziman, J. (2000). *Real science: what it is, and what it means*. Cambridge: Cambridge University Press.