

# MINDING OUR PS AND QS – ISSUES OF PROPERTY, PROVENANCE, QUANTITY AND QUALITY IN INSTITUTIONAL REPOSITORIES

**Bruce White**

Massey University, New Zealand

[b.d.white@massey.ac.nz](mailto:b.d.white@massey.ac.nz)

## **Abstract**

The development of institutional repositories has opened the path to the mass availability of peer-reviewed scholarly information and the extension of information democracy to the academic domain. A secondary space of free-to-all documents has begun to parallel the hitherto-closed world of journal publishing and many publishers have consented to the inclusion of copyrighted documents in digital repositories, although frequently specifying that a version other than the formally-published one be used. This paper will conceptually examine the complex interplay of rights, permissions and versions between publishers and repositories, focussing on the New Zealand situation and the challenges faced by university repositories in recruiting high-quality peer-reviewed documents for the open access domain. A brief statistical snapshot of the appearance of material from significant publishers in repositories will be used to gauge the progress that has been made towards broadening information availability. The paper will also look at the importance of harvesting and dissemination, in particular the role of Google Scholar in bringing research information within reach of ordinary internet users. The importance of accuracy, authority, provenance and transparency in the presentation of research-based information and the important role that librarians can and should play in optimising the open research discovery experience will be emphasised.

## **Introduction**

The digital institutional repository of theses, reports and published documents has become sufficiently well established within academic and research institutions to warrant a consideration of its progress and potentialities. The Institutional Repository movement has sought to stem the expropriation by private interests of valuable intellectual property originating within the research community as a means of both improving the research infrastructure and controlling public sector costs, but a major British government report [United Kingdom Select Committee on Science and Technology, 2004] concluded that individual academics were content to ignore issues surrounding the ownership of intellectual property. In a report to the Joint Information Systems Committee Heery and Anderson [2005] also noted the existence of important cultural barriers and a lack of communication and understanding between repositories and scholars. While voluntary archiving has had a disappointing uptake [Xia & Sun, 2007b] there is some evidence that compelling authors to contribute material to repositories may be more effective [Cochrane & Callan, 2007; Sale, 2006]. An understandable emphasis on recruiting material for repositories in the start-up phase may also have been accompanied by a relative tolerance on quality issues and it is notable that in her proposed indicators of success Westell [2006] did not include quality of either documents or metadata as critical factors. This paper will examine the complex web of rights and permissions surrounding the archiving of published journal articles and the importance of creating repositories of high quality and documents of clear identity and provenance.

Libraries have traditionally accorded a privileged status to publications as opposed to other forms of documentation and communication – as their very name suggests libraries have collected *books* rather than unpublished writings and they have taken seriously the distinction between those writings that have appealed to a publisher as being worthy of placing between covers and issuing to the market and those that have not. To an extent this has been a matter of convenience – printed books are a relatively cheap artefact to store and lend – but the fact of publication also carries with it an expectation of worth and quality that might be assumed to be absent from a work that has been rejected for publication. (Anyone who doubts this might try to interest people in reading the manuscript of an unpublished novel.) Where non-printed materials have been collected and stored this has been because of their particular value or interest, generally historical. Non-printed material is not normally collected *as a matter of course* whereas libraries will choose to purchase certain types of print material on the basis of subject or genre alone.

### **The Scholarly Journal System**

This is true of all libraries but in academic libraries judgements of value on the basis of publication are particularly prevalent and fine-grained discriminations on the basis of journal title, publisher, author's host institution and so forth are the librarian's stock-in-trade. Unqualified to make direct judgements about the content of most of the works we deal with we are nonetheless quite able to use proxy methods by considering issues of provenance and context, directing students away from random websites and blogs and towards encyclopedias, scholarly books and peer-reviewed journals. This is possible because these hierarchies of value are built into the structure of the academic literature itself and the external signals are generally consistent with a broadly shared consensus of values among the scholarly community. Looking at the journal literature as the most significant and highly structured instance of academic publishing we find a sophisticated system of editorial control and peer-reviewing that mirrors the structure and in fact recruits the personnel of the specific discipline with which each journal is concerned. Journals and academic communities are so closely related to one another that the journal can be regarded as an outward expression of the community while for an individual being denied access to the journal can effectively deny them membership of the community.

Editorial oversight and peer review systems are a means by which the subject-based community exercises control and discipline over its members, ensuring that in the production of new knowledge or ideas they continue to work with proven methods and within accepted paradigms. Although much of this work is unpaid, journal editing and publishing (if done to a high standard) do carry significant costs which, combined with low gross demand for most specialist titles, can result in surprisingly high per title subscription prices. Because access to the journals is critically important to the communities they serve academic libraries have seen themselves as having little choice but to pay these prices and the result has been high profit margins and the price spiral known as the "serials crisis". One of the major drivers behind this process is copyright, the requirement that authors transfer intellectual property rights to the journal publishers as a condition of having their work accepted for publication, and it has been long been apparent that work being produced within universities was being given away and then sold back to them at inflated prices.

A further and equally unfortunate outcome of the journal system has been that access to scholarly work is generally restricted to members of the academic community. High prices and low circulations means that specialised journals are held only by the libraries of institutions working within the relevant discipline and although many university libraries are at least in theory open to the public for browsing purposes there are formal and informal barriers to access that only the most confident, knowledgeable and determined non-affiliate is likely to penetrate. Again in theory, interlibrary loan systems make all publications available to all citizens of modern democracies through their public libraries but lack of access to the bibliographic infrastructure of indexes and abstracts needed to negotiate the literature has made information democracy a largely unrealised ideal. Indeed the Select Committee on Science and Technology [2004] concluded that "we are not convinced that journal articles are consistently available to members of the public through public libraries"

## **Electronic Developments**

The advent of the internet in the early 1990s held out the promise that libraries and scholars could be released from the grip of this hegemony by mobilising the free scholarly labour being given to the publishers and reinventing the journal system through a network of free internet journals originating within and controlled by the academy itself. As well as breaking the link between authors and publishers this movement also promised to strengthen links within the academic community and to make academic publications widely available through a revolution in dissemination. That this early initiative had only a very limited impact was probably down to two factors. The notion that academic labour alone could sustain a journal proved to be flawed and it was discovered that some kind of income stream was necessary to sustain quality academic publishing. Secondly, authors sought the recognition of having their work appear in established journals of high prestige and may even have actively avoided the low status accorded to what were perceived as shoestring operations.

A somewhat later development was the open access journal which was a high-quality peer-reviewed publication available to all users of the Internet. Open access publishers addressed the question of income flow by reversing the normal requirement for payment, which was transferred from the consumers (subscribing institutions and readers) to producers (authors or their host institutions who paid the costs of editing, reviewing and publishing). From the point of view of both library budgets and information democracy this has been a positive development but at the cost of restricting access to publishing outlets for authors unable to come up with the money. A doubt might also exist as to whether a journal can be said to publish the "best" articles in its field if the selection process is in part financially determined. Nevertheless because it has secured an income stream the open access movement continues to thrive and grow and several journals of this kind have established themselves as major titles alongside those put out by the major conventional publishers. Because the financial transaction occurs prior to publication of the article rather than through on-sale of the work copyright provisions relating to open access journals allow pretty much open use and distribution subject to proper attribution.

The second half of the 1990s saw the transition from purely print publishing to a print/electronic hybrid on the part of the major academic publishers. Articles were produced and submitted to journals electronically and then converted into a printed journal issue and a set of PDF files identical in appearance to the articles in the printed journal. In theory an electronic iteration of a work can be quite different in appearance from its print "original" but in practice print has continued to retain its status as the standard version from which others are perceived as varying. There are some very real practical benefits to maintaining the appearance of print in digital formats such as the ability to cite page numbers but one of the major factors is undoubtedly psychological, the branding of the article as part of a printed journal of known provenance, extent and status.

## **Journal Article Typology**

For many years now most academic journal articles have been "born digital", written and edited on the computer and generally passing through numerous iterations and refinements on their way to becoming printed artefacts. Although we use the terms "pre-print" and "post-print" to refer to clearly defined stages through which an article passes these distinctions are less absolute in reality. A "pre-print" is commonly defined as a version of the paper before it has been through the peer-review process whilst a "post-print" incorporates the changes made as a result of that process but is not necessarily the "printed" version. Normally the final Microsoft Word document sent by the author to the journal editor is regarded as a post-print even though it may differ in some details from the final published article. The degree of difference between pre-prints and post-prints can vary greatly; sometimes they are relatively minor and even difficult to detect while in other cases the post-print represents a substantial revision of the originally-submitted article to the extent that it constitutes a quite separate entity. Similarly there can be differences between different post-print versions of the same

article which although relatively subtle are nonetheless important. To determine the effects of the copy-editing process Wates and Campbell [2007] examined the differences between authors' final versions of articles and publishers' versions and concluded that although "none of these [changes] materially altered the conclusions of an article ... they did produce a more consistent and accurate article of record." This finding was confirmed by Goodman, Dowson and Yaremchuk [2007]

In this context the notion of an "article of record" is perhaps the critical one and print has the advantage here in its absolute resistance to retrospective alteration which is substantially transferred to its electronic analogue, the so-called "publisher PDF". The significant differences that Wates and Campbell detected between authors' post-prints and published articles are by their very nature hidden unless the copy under examination is compared to the "genuine article". As long as this is defined as a particular and unique arrangement of words and images on the (numbered) page two ontological facts are established almost beyond doubt – a) this article by these authors was published in this particular issue of this journal on these pages and b) this *is* the article that was published etc etc. Any other version has an authenticity deficit on both these counts – the article may not have been published in the specific journal as claimed and, if it was, the version under examination may differ from the published article.

This question is of more than purely theoretical importance. When the first databases containing full-text were introduced in the 1980s and 1990s the articles contained were either pure text versions or html web pages containing all the words and images but without the precise format of the print article. Many librarians were disappointed to discover that having read the text of an article researchers frequently wished to obtain a photocopy of the original on interlibrary loan. It was only when the PDF format became widely available that the online mode became an acceptable, indeed preferred, form in which to read journals. The same phenomenon can be observed in download statistics of electronic journals where those for html versions are invariably much lower than those for PDFs. Initially this was dismissed as pure conservatism, a question of "look and feel" that would pass as users became accustomed to new formats but in fact the preference has proven to be deeply ingrained and is arguably closely related to the issues of authenticity and accuracy outlined above.

### **Institutional Repositories**

Scholars and libraries were quick to see the benefits of the electronic journal and indeed the advent of the so-called "big deal" in the late 1990s saw massive increases in the availability and affordability of journal titles, but unfortunately no decrease in gross costs as libraries saw an increasing percentage of their budgets swallowed up by offers they couldn't refuse; they (and their clients) had never had it so good but had it never cost them so much. The ongoing concern that the content that constituted the actual value of the journals originated within the academy but was being given away through copyright assignment and purchased back through subscriptions lead to the development of the digital repository, holding and making available digital copies of published articles, as a means of circumventing this closed system. An early subject-based repository, the American Institute of Physics *arXiv*, held pre-prints rather than post-prints but latterly the trend has been towards institutional repositories holding post-prints of published articles. The obvious barrier to this process has been the fact that copyright of published articles has been assigned to publishers and despite Stevan Harnad's contention that relatively minor changes and enhancements to the published version would create a new copyright-free entity [Harnad, 1999] this has proven not to be the case with authors overwhelmingly continuing to assign copyright in order to have their work published and publishers continuing to assert ownership of the articles.

For obvious reasons publishers have not seen institutionally-based repositories as posing the same degree of threat to their subscriptions as subject-based ones and the majority of them have allowed post-prints of articles to be stored and made generally available through institutional repositories. The catch, however, and it is a big one, is that the major academic publishers, with one or two notable exceptions, have not allowed their own PDF print-analogue version to be stored but rather the author's final version, post refereeing but pre-

copy-editing and not of course showing the pagination and branding of the “genuine article”. Usually they insist that their own version be linked to through a url or a digital object identifier and that the author’s version be used in the repository.

If the institutional repository is to play a role in increasing the availability of peer-reviewed literature both within the academy and beyond that to the wider community, in other words if it is to have an impact on the cost to universities of purchasing journals and to enhance information democracy, then it will do so through making available material that would otherwise have been restricted to subscribers only. In simple terms, if repositories were making a substantial impact we would expect to find significant numbers of author PDFs giving unrestricted access to all. A detailed survey of the availability of documents through repositories poses formidable difficulties in terms of defining exactly what a repository is and how, and in fact whether, it differs from departmental or personal research pages on university servers, all of which can be detected when searching for documents on the Internet, and it is beyond the scope of this paper to attempt this task. However a Google Scholar search for 100 documents published in 2004 and previously identified through Web of Science found four author’ post-prints held in repositories as opposed to a total of twelve publisher versions, suggesting that the publishers’ permission to include author post-prints in repositories has not been widely taken up. Seven documents were found that appeared to be author pre-prints. It is not possible to draw conclusions as to the greater numbers of publishers’ versions available without knowing whether permissions had been sought and granted for these but it is fair to say that departmental and personal research pages appear less likely to comply with copyright requirements than formally constituted repositories.

**Table 1: Fulltext access to a sample of 100 journal articles from Google Scholar (December 2007)**

Fulltext Restricted*	72
Fulltext Open Access	8
Repository Preprint	7
Repository Author Postprint	4
Repository Publisher Postprint	12
*Refers to publishers’ websites or other paid for versions of articles	

As the sample size was exactly 100 raw figures represent percentages.

An examination of the numbers of fulltext journal articles held in New Zealand university repositories showed low numbers in relation to the total amount of publishing.

**Table 2: Journal articles in New Zealand university digital repositories 14 March 2008**

University	Journal articles in repository
Massey	98
Auckland*	71
Waikato	236
Victoria*	20
Canterbury	256
Otago	29
AUT	0
Lincoln	8
*Estimate	718

Extracting this data proved to be difficult as some repositories do not organise their contents by document type but the figures are unlikely to understate the holdings. Furthermore as they

relate to numbers of individual records in the repositories, not all of which have documents attached, the true number of journal articles being made available through this means is somewhat smaller than the figures indicate. That these figures represent only a small proportion of total journal article publishing is evident from the fact that a total of 4,386 articles from the same universities published in 2007 appear in the *Web of Science* database alone while the 2006 Performance-Based Research Fund assessment reported a total of 32,209 journal articles from New Zealand tertiary institutions for the period 2000-2005 [New Zealand Tertiary Education Commission, 2006].

Submission of material to New Zealand repositories is voluntary and these statistics form an interesting comparison to those from the Queensland University of Technology (QUT) which states in its Manual of Policies and Procedures that, subject to certain exemptions, "material which represents the total publicly available research and scholarly output of the University is to be located in the University's digital or " E print " repository" [Cochrane & Callan, 2007; Queensland University of Technology, 2006]. In total the QUT ePrints Archive contains records for 5,484 journal articles; of these 676 were published in 2005 from which an analysis of 200 records revealed that 59% were publicly available full text documents. A further 4% were full documents that were not publicly available while 37% of records did not have attached documents.

**Table 3: Sample of 2005 published journal articles from QUT ePrints Archive<sup>1</sup>**

Author PDF	102	51%
Publisher PDF	16	8%
Bibliographic record only	74	37%
Restricted	8	4%

200

If we look at only those articles for which fulltext is available we find that 14% are publisher PDFs which is not dissimilar from the figure from our Google sample in Table 1.

**Table 4: Sample of 2005 fulltext journal articles from QUT ePrints Archive**

Author PDF	102	86%
Publisher PDF	16	14%

118

Full text documents in the QUT repository all have a cover page that includes a full citation which means that the critical identifying metadata are embedded within the document which is in no danger of becoming an 'orphan'.

An examination of 131 journal articles held in a discrete collection in a New Zealand university repository reveals a somewhat different pattern of document types and provenances with a higher proportion of publisher PDFs (22%) and a significant number of preprints and publishers' proofs..

**Table 5: Analysis of fulltext articles from a New Zealand university repository A**

---

<sup>1</sup> It should be noted that Table 3 is at considerable variance to the 96% rate of full text availability for QUT found by Xia and Sun [2007a] which may have referred to a wider range of document types. A further check of 100 articles from 2006 and 2007 also revealed that 37% of them did not have attached fulltext

Author postprint PDF without metadata	55	42%
Author postprint PDF with metadata	4	3%
Preprint	10	8%
Proof copy	11	8%
Publisher PDF	29	22%
Scan of published article	3	2%
Word postprint, no metadata	18	14%
Word postprint, some metadata	1	1%

131

A discrete collection from a second repository reveals a quite different pattern.

**Table 6: Analysis of fulltext articles from a New Zealand university repository B**

Author PDF without metadata	22	36%
Author PDF with metadata	1	2%
Publisher PDF or scan	38	62%

61

The present numbers of fulltext journal articles in New Zealand repositories are too small for any meaningful pattern to emerge and current practice is probably determined by a small number of authors who are keen contributors on the basis of either belief or perceived self-interest. Comparing the New Zealand experience with that of QUT it is clear that voluntary submission results in a very small capture of the potential target and that without widespread education and support in addition to coercion we are unlikely to create a substantial collection of high quality documents to add value to existing bibliographic structures and to advance the cause of information democracy. Recent moves by the National Library of New Zealand in the direction of mandating submission to repositories can be expected to succeed only if accompanied by the funding of repositories to work with academics in the harvesting of research.

### **Discussion and Implications**

The difficulties experienced in obtaining author post-prints are at first puzzling given that it is this area in which the greatest promise lies for institutional repositories to make a difference. Barwick [2007] noted the difficulties of obtaining author versions of articles and anecdotal evidence abounds that authors are either unable or unwilling to provide the final version of their article sent to the publisher post-refereeing. Part of the problem is undoubtedly lack of awareness on the part of authors of the significance of this digital object and the importance of retaining it but probably as important as this may be pure reluctance on the part of authors to have a secondary version of the article in circulation once the published version has appeared. A minority of authors are probably active believers in open access publishing and will seek out open access journals and pay to have their work published in this mode – it may even be the fact that high profile authors with the ability to attract research funding are more likely to use this mode that accounts in part for the success of many open access titles. However the majority are probably at best agnostic on the subject and are unlikely to actively seek to increase the availability of their work beyond the scope of traditional journal-subscribing academic and research institutions.

To understand why this should be we need to consider briefly the privileged position of the scholarly journal in the academic community and in the consciousness of its members.

Having had an article published in a journal of quality and significance the researcher is wholly satisfied – they have placed their work on record within the community of interest, they are able to satisfy the requirement of their employing institution for published research outputs and beyond that they have few concerns. To the author whose work has been accepted for publication by a respected journal the printed copy of the article as represented by the publisher PDF amounts to a certification of its worth and to place a less prestigious version into circulation represents a potential threat to this status. If repositories are to mount significant numbers of author post-prints they need their cooperation in saving and supplying final versions and this is unlikely if they are poorly motivated. In these circumstances it is also even less likely that they will choose to self-archive. On the other hand when publishers allow their own versions to be held in repositories authors are enthusiastic about their inclusion.

Whilst there is evidence that the wider availability of an article afforded by open access increases its rate of citation [Antelman, 2004; Eysenbach, 2006; Harnad & Brody, 2004] the logic of this is not immediately obvious to many authors who reasonably enough expect that those working within the field and therefore most likely to cite them will have subscribed access and will keep themselves abreast of the literature. This may be partly a misconception by authors who see their specialisation as more bounded than it in fact is and who do not take account of the possibility of citations from outside their own immediate field of research. Beyond that a “public good” appeal to the value of disseminating research beyond the academy is not necessarily convincing to the scholar who may feel that broader exposure of their work carries some distinct disadvantages in terms of unwanted approaches from non-affiliates and even liability issues. The inclusion of an email address with an article is a helpful means of facilitating communication between scholars but may take on a quite different connotation when the information is widely disseminated to curious and even potentially hostile members of the public. At the very least the importance of communicating research findings to the public has a much lower priority than communicating with peer networks [Kennan, 2007]. Globalisation has brought scholars closer together and has ameliorated in particular the isolation felt by scholars in the south-west Pacific and apart from those who choose to adopt a high public profile, or those with a strong ethical or political commitment to the public archiving of research, the scholarly community is at best only mildly sympathetic to the difficulties experienced by the public in gaining access to research.

Once articles are placed within a repository they become available to a variety of search engines including those specifically covering repositories such as OALster and also to broader sources such as Google Scholar which is used by a far larger number of researchers and is likely to be the means by which most of them come into contact with the contents of repositories. By covering ‘everything’ Google Scholar outperforms other sources in quantity of results and is always going to be a place where searchers go as a matter of course. While considerable energy has gone into designing repositories with a pleasing appearance, layout and searchability it is unlikely that more than a tiny fraction of users will come to the repository itself rather than to one of its constituent documents through a search engine. Despite its shortcomings and tendency towards bibliographic bloopers through the sheer extent of its coverage and its full-text search capability Google Scholar now stands as a rival to traditional databases and is the tool of preference for a growing number of scholars [White, 2006]. Covering as it does most repositories of significance its arrival on the scene in 2004 may in fact be seen as providing a considerable boost to their accessibility and functionality. Even within the scholarly domain the percentage of potential users of scholarly information who would even be aware of such specialised search engines as OALster or KRIS (Kiw Research Information Service) is probably tiny compared to the users of Google and Google Scholar and the primary advantage of such utilities is as a reliable source of structured records for the large popular engines. The non-adoption by Google Scholar of open metadata standards is a significant disappointment in this regard.

Writing of Google Scholar searches of repository material Markland [2006] noted the “the difficulty of identifying the exact status of the document presented in the repository”; if most searchers come to repository records and documents from “outside” they may have a very limited awareness of the context within which these are situated. Many repository records consist of metadata with attached PDFs but the searcher may well harvest the PDF and move on, taking little notice of the metadata. Unless the document itself proclaims its full identity it

is in danger of becoming a decontextualised orphan sitting in a folder on a computer. Tables 5 and 6 above (and almost any examination of items held in repositories with some notable exceptions) reveals a number of papers that have a title and author but no other bibliographic data and no indication of status – they could be pre-prints or post-prints of journal articles, they could be conference papers, they could be working or discussion papers. As long as they remain safely within the repository attached to the metadata these questions are easily answered, but as soon as they are downloaded, unless a conscious effort is made by the user to attach them to a citation, they are soon unidentifiable, a headache for researchers, research assistants or librarians and less likely to be cited. As an absolute minimum each article, be it pre-print, author post-print or publisher post-print, requires a cover page or initial paragraph containing the full citation and the link to the article on the publisher's website or the digital-object-identifier if this is a condition for inclusion in the repository. Further information identifying the repository itself and also which of the authors is linked to the host institution would be useful as well.

Another issue that repositories need to address is the inclusion of harvestable bibliographic records with no actual documents attached to them or attached to passworded, closed-access documents. Whilst it may be desirable for institutions to record and highlight their publications the primary function of a repository is to store and deliver documents. The world is not short of metadata and most articles already have records in at least one of the major public databases such as PubMed, Ingenta or CSA Illumina. Searchers are looking for documents and the unnecessary proliferation of these records merely serves to clog up the system and slow down the process.

### Future Prospects

Repositories are still in their very early years and the full effects of such developments as mandatory submission are yet to be seen. In this part of the world few repositories have yet tackled the task of storing the journal literature and yet it is in this area that much of their promise has been seen to lie. Much progress is being made on the technical front but formidable cultural and behavioural barriers still exist. The active engagement of librarians in the process is critical because of our relatively greater exposure to the research process and questions of information accessibility and quality. Some of the early bright hopes of the open access movement may remain unrealised and no immediate solution to the serials crisis is yet in sight. Information democracy has already been immensely enhanced by many features of the new order but also remains an elusive goal with non-affiliated researchers still locked out from much of the peer-reviewed literature and having to make do with much secondary material. It is important, however, that these initial goals be kept clearly in sight or the repository will become a storehouse of theses and dissertations, conference papers, working papers and various odds and ends but lacking substantial peer-reviewed scholarly journal content. While national mandating may represent the best hope for the repository movement organising academics has been likened to herding cats and we should not under-estimate the cultural, intellectual and legal barriers that are likely to be faced.

### References

- Antelman, K. (2004). Do open-access articles have a greater research impact. *College & Research Libraries*, 65(5), 372-382.
- Barwick, J. (2007). Building an institutional repository at Loughborough University: some experiences. *Program-Electronic Library and Information Systems*, 41(2), 113-123.
- Cochrane, T., & Callan, P. (2007). Making a difference: implementing the eprints mandate at QUT. *OCLC Systems & Services*, 23(3), 262-268.
- Eysenbach, G. (2006). Citation advantage of open access articles. *PLoS Biology*, 4(5), e157.
- Goodman, D., Dowson, S., & Yaremchuk, J. (2007). Open access and accuracy: author-archived manuscripts vs. published articles. *Learned Publishing*, 20(3), 203-215.
- Harnad, S. (1999). Free at last: The future of peer-reviewed journals. *D-Lib Magazine*, 5(12).

- Harnad, S., & Brody, T. (2004). Comparing the impact of open access (OA) vs. non-OA articles in the same journals. *D-Lib Magazine*, 10(6).
- Heery, R., & Anderson, S. (2005). *Digital repositories review*. Retrieved 17 March 2008, from [http://www.jisc.ac.uk/uploaded\\_documents/digital-repositories-review-2005.pdf](http://www.jisc.ac.uk/uploaded_documents/digital-repositories-review-2005.pdf).
- Kennan, M. A. (2007). Academic authors, scholarly publishing, and open access in Australia. *Learned Publishing*, 20(2), 138-146.
- Markland, M. (2006). Institutional repositories in the UK: What can the Google user find there? *Journal of Librarianship and Information Science*, 38(4), 221-228.
- New Zealand Tertiary Education Commission. (2006). *Performance-Based Research Fund: Evaluating research excellence. The 2006 assessment*. Retrieved 16 March 2008, from <http://www.tec.govt.nz/upload/downloads/pbrf-quality-evaluation-report-25-May.pdf>.
- Queensland University of Technology. (2006). Manual of policies and procedures: F/1.3 E-print repository for research output at QUT. Retrieved 17 March 2008, from [http://www.mopp.qut.edu.au/F/F\\_01\\_03.jsp](http://www.mopp.qut.edu.au/F/F_01_03.jsp)
- Sale, A. (2006). The acquisition of open access research articles. *First Monday*, 11(10).
- United Kingdom Select Committee on Science and Technology. (2004). *Science and Technology - Tenth Report, Session 2003-04*. Retrieved 14 March 2008, from <http://www.parliament.the-stationery-office.com/pa/cm200304/cmselect/cmsctech/399/39902.htm>.
- Wates, E., & Campbell, R. (2007). Author's version vs. publisher's version: an analysis of the copy-editing function. *Learned Publishing*, 20(2), 121-129.
- Westell, M. (2006). Institutional repositories: proposed indicators of success. *Library Hi Tech*, 24(2), 211-226.
- White, B. (2006). Examining the claims of Google Scholar as a serious information source. *New Zealand Library & Information Management Journal*, 50(1), 11-24.
- Xia, J. F., & Sun, L. (2007a). Assessment of self-archiving in institutional repositories: Depositorship and full-text availability. *Serials Review*, 33(1), 14-21.
- Xia, J. F., & Sun, L. (2007b). Factors to assess self-archiving in institutional repositories. *Serials Review*, 33(2), 73-80.

# Minding our ps and qs: Issues of property, provenance, quantity and quality in institutional repositories

White, Bruce

2008

---

<http://hdl.handle.net/10179/7637>

22/09/2020 - Downloaded from MASSEY RESEARCH ONLINE