

ISSUES OF PROPERTY, PROVENANCE, QUANTITY AND QUALITY IN
INSTITUTIONAL REPOSITORIES

MINDING OUR PS AND QS

Bruce White

Massey University Library

New Zealand

P's and Q's, *n.*

DRAFT REVISION Mar. 2008

1612 S. ROWLANDS *Knaue of Harts* (Hunterian Club) 20 Bring in a quart of Maligo, right true: And looke, you Rogue, that it be Pee and Kew. **1876** T. M. BOUND *Hereford. & Shropsh. Provinc. in Eng. Dial. Dict.* at P, To be P and Q, to be of prime quality.

3. *one's P's and Q's*: one's alphabet, one's 'A.B.C.'; (in extended use) knowledge or judgement, 'one's way about'. Also: good manners, etiquette (cf. sense 1).

Connotations of good manners and correct behaviour are often present even when they are not explicitly referred to.

1763 C. CHURCHILL *Poems* IV. 351 On all occasions next the chair
He stands for service of the Mayor, And to instruct him how to
use His A's and B's, and P's and Q's. ?**1804** A. PASQUIN

Outline

- Libraries and publications
- The scholarly journal system
- Electronic developments
- Journal article typology
- Repositories
- A look at some repositories
- Implications



Libraries and the publication process

- Libraries, as opposed to archives, are collections of “finished products”
- Academic (and other) libraries rely on editors and publishers to exercise gateway and quality control functions
- Publishers own information through copyright



Academic libraries

- In academic libraries judgements of value on the basis of publication are particularly prevalent
- Hierarchies of value are built into the structure of the academic literature itself
- Librarians steer clients towards formally-published sources



The scholarly journal system

- The scholarly journal is an expression of its discipline community
- The discipline community exercises control over its members through the journal
- Lack of access to the journal can exclude individuals from participation in the discipline – gated community
- Costs of quality control and pressure to subscribe have resulted in high per title prices



Electronic developments

- The Internet promises the broadening of access to the scholarly literature
- The Internet Journal – low cost handicraft model
- The pre-print server
- The PDF, the print/electronic hybrid and the persistence of print “look and feel”
- The “big deal” maintains the profits of the big publishers while lowering per title prices
- The open access journal moves costs away from consumers
- The digital repository offers an alternative source of supply



Journal article typology

- Pre-print – before peer-review
- Post-print – after peer review but before copy-editing and publishing
- Post-prints may or may not include cover pages or other embedded identifiers – “orphan papers”
- Publisher’s version – print analogue “copy of record”
- Different versions lead to issues of quality, provenance and authenticity



Institutional repositories

- Repositories based on institution pose little threat to publishers
- Offer some hope of controlling journal prices and enabling information democracy
- Publishers generally allow inclusion of post-print but not of the print analogue “published version”
- Voluntary take-up by authors disappointing even if they are sympathetic



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

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
	718

*Estimate

Sample of 2005-published journal articles from QUT ePrints Archive

Author PDF	102	51%
Publisher PDF	16	8%
Bibliographic record only	74	37%
Restricted	8	4%
	200	

NB Submission of research outputs to the archive at QUT is mandatory. All PDFs have cover pages

Analysis of fulltext articles from a New Zealand university repository

Author postprint PDF without cover page	55	42%
Author postprint PDF with cover page	4	3%
Preprint	10	8%
Proof copy	11	8%
Publisher PDF	29	22%
Scan of published article	3	2%
Word postprint, no identifying metadata	18	14%
Word postprint, some metadata	1	1%
	131	

Discussion

- While most authors are sympathetic to open-access they don't necessarily see a personal benefit
- Authors tend to be concerned with their own discipline community
- Author attitudes to versions other than the published one create a barrier to submission
- Mandatory submission may be the answer but needs to be accompanied by huge support

Discussion cont...

- Users of documents are more likely to come to the repository through a search engine, especially Google Scholar
- Repositories cannot rely on context to provide publishing metadata
- Risk of creating “orphan documents”
- Repositories need to adopt consistent cover page standards for embedding publication metadata into documents
- Bibliographic records without documents are of dubious worth

Conclusions

- Repositories have a long way to go in addressing the serials crisis
- They are an excellent publication channel for grey literature, working papers etc, but ...
- they will fulfil their potential only if substantial amounts of peer-reviewed literature are included