

USING INSTITUTIONAL REPOSITORIES SOME RECURRING CONCERNS AND CLARIFICATIONS

Circulation: PUBLIC
Bill Hubbard
University of Nottingham

The rapid growth of institutional repositories and the possibilities they bring for enhanced dissemination and the use of published work has brought with it some uncertainties and misunderstandings. Listed below are some of the more common misunderstandings and some clarification of the issues.

** That repositories should be used *instead* of publishing in journals*

On the contrary, authors should continue to publish in journals, whether traditional or Open Access. Repository use is *supplementary* to journal publication.

** That repository material will bypass peer-review.*

No - articles undergo peer-review through the normal journal publication process before mounting in the repository. Where some subject-disciplines use non-refereed pre-prints or working papers to develop their ideas, then repositories can support the practice for that discipline - but it is clearly labelled as such.

** That if articles are easily available, then plagiarism will be made easier*

In fact, plagiarism is diminished as a problem. It is far easier to detect if the original, date-stamped material is freely accessible to all, rather than being hidden in an obscure journal.

** That authors will be forced to give IPR to the University*

The use of institutional repositories is tied into a drive to give authors *more* control over their work and the way they want to use it, not less. This means Universities clarifying their IPR policies in favour of authors, and for authors to retain copyright on publication, rather than having to assign it to publishers and accept restrictions on re-using their materials in teaching, sharing with colleagues, etc

** That publishers do not allow material to be archived in this way*

Recently several big publishers - Springer, Elsevier and others, have adopted "repository-friendly" policies. Together with those publishers that have always allowed authors to mount their work on-line, such moves have meant that the standard database of publisher's policies (SHERPA/RoMEO list) now shows that 64% allow post-print archiving - a figure which rises to over 90% when considered by the number of journals, rather than by publishers.

** That authors of books and reference works will lose royalties*

None of this applies to work from which authors expect to make money. Authors writing text-books or articles on commission quite rightly do not want to give away their work. Institutional repository use is aimed at research information that authors want to be spread and read as widely as possible.

* *That commercially sensitive material or patentable inventions will be given away*
No-one is suggesting that such material be given away. If material is confidential or sensitive, then it is not put into a repository, in the same way that it would not be published in a journal. Repository use is for material that authors want to see disseminated, cited and made public.

* *That this is just aimed at computer scientists*

This is not just for any specialised discipline - any researcher in any field can use their institutional repository. SHERPA repositories currently include eprints on Romanticism, botany, sociology, pathology and Burmese linguistics.

* *What about version-control? Will there be one version in a repository and a different published version?*

Most deposited versions are the author's final version - the version after peer-review, after revision and checking - that is finally sent to the publisher for printing. Some publishers allow the use of the final PDF file which contains their layout and style of font - in which case the versions are identical. Otherwise, while the style of the font might be different, the text of the repository version is identical except for whatever minor copy-editing is done by the publisher after it leaves the author's hands.

* *That repository use will kill journals.*

The available evidence suggests otherwise. The Physics community has been using a subject-based archive - arXiv - to hold their papers for over thirteen years, while also sending them to traditional journals: and yet Physics journals continue to be sold. arXiv now holds over 300,000 papers and takes in about 1/3 of the world's physics research. A recent survey asked the society publishers *American Physical Society* and the *Institute of Physics* the same question - "Has the use of arXiv made you lose subscriptions?"

Both replied "No."

The evidence shows that repository use does not kill journals and that repositories can exist as a stable, supplementary but highly advantageous form of dissemination.

* *The "Free Rider Problem"*

This is essentially a commercial concern: if research results are available for free, then some companies will be able to access and use the articles for their own commercial advantage without paying for it. This has been voiced as a concern by some publishers, who feel that commercial companies should pay the publishers for access. It has been observed as ironic that some publishers feel it is unfair if commercial companies make profits out of articles that are given to them for free.

Some publishers get a lot of their income from commercial subscriptions and it is of course completely natural that publishers be concerned about the stability of their business model and the level of their profit margins. However, companies acting as "free-riders" do not financially damage the academic or the research institution as producers of the research: indeed it could be argued that the greater availability and use of research could foster closer ties between industry and university researchers. Again, more rapid and wider dissemination of the latest research can only promote innovation, development and industrial growth.

** That this is an imposition on academic authors.*

Mounting an article on a repository does not require a heavy additional workload for authors. Given an electronic copy of the final text - the one that is finally sent to the publishers - the process is fairly straight forward. After registration with an institute's service the task basically consists of filling in an on-line form and hitting the "submit" button. Once an author has gone through the process for the first time, then it should take less than 10 minutes per eprint. Measured against the time it takes to get the funding, do the research and write the article, 10 minutes work is a minimal cost for the increased exposure and citation it gains.

** "I have already got my material on my web-site - so I don't need to put it in a repository"*

Institutional repositories are based on a standard system from the Open Archives Initiative (OAI). Because the metadata - effectively the catalogue description - in OAI-compliant repositories is "harvested" by OAI search services, all such repositories can be searched through a single point. Since the only results of such a search come from academic repositories, an eprint is far easier to find and far more prominent than if it were accessed through an individual's non-OAI website. Even a normal Google or Google Scholar search favours OAI-repository material and normally ranks it higher than an individual's own website.

More than this, repositories are working to help material be preserved in the longer term. This brings the advantage that when academics move on, or their personal website changes, their eprints in a repository - and the links - will remain stable, readable and accessible.

** A subject-based repository would be far better - why should eprints on chemistry be in the same repository as eprints on linguistics?*

To someone searching for a particular eprint, the fact it is on an institutional repository is irrelevant. The OAI system ties all the subject-based and institutional repositories together into one massive virtual archive. A search service will just find those articles that match the search, present the abstract of the eprint and the reader can view the full-text if required. The access point or the route by which the eprint is *viewed*, is independent of the place it is *stored*.

** Are institutional repositories in competition with subject-based repositories or open-access journals?*

Institutional repositories do not clash with open-access journals; or with subject-based repositories like arXiv; or the economists repository, rePEc; or with personal web-pages; or with departmental web-pages, or wherever else an eprint can be mounted. One of the strengths of the Open Access movement is that different projects are not in competition with each other. When dissemination is the aim, then the more players the better.

SHERPA Project Document

USING INSTITUTIONAL REPOSITORIES
- SOME RECURRING CONCERNS AND CLARIFICATIONS

Bill Hubbard
University of Nottingham
February 2005