Researcher Attitudes to Data Sharing: Cultural Change Requires Better Motivations

Philippa Broadley¹, Kerry Raymond²

¹Queensland University of Technology, Brisbane, Australia, p.broadley@qut.edu.au
²Queensland University of Technology, Brisbane, Australia, k.raymond@qut.edu.au

INTRODUCTION

While undertaking the ANDS RDA Gold Standard Record Exemplars project, research data sharing was discussed with many QUT researchers. Our experiences provided rich insight into researcher attitudes towards their data and the sharing of such data. Generally, we found traditional altruistic motivations for research data sharing did not inspire researchers, but an explanation of the more achievement-oriented benefits were more compelling.

BENEFITS OF DATA SHARING

The traditional altruistic benefits of sharing research data with other researchers include:

- Making public the outcomes of research funded by public monies
- Other researchers do not needlessly repeat existing research (saving time and money)
- Building on the research of others (“standing on the shoulders of giants”)  
- Using data to answer new research questions

... all of which seek to further scientific discovery for the benefit of humanity.

QUT supports the sharing of research data, saying “Research data should generally be made available, via open access, for use by other researchers unless a specific and valid reason exists for not doing so.”[1]

However, in approaching QUT researchers in relation to the Gold Standard project, many researchers felt unable to make their research data available, usually citing concerns in relation to the ethics approval of the project that collected the data or because of constraints imposed by industry partners in the research project.

Given that QUT is a “university for the real world”, many research activities are subject to ethics approval and third party contracts. Consequently, the Gold Standard project had to seek out those projects not so constrained, but those researchers often expressed doubts about the value of data sharing.

CONCERNS ABOUT DATA SHARING

Generally, researchers did not find the aforementioned altruistic benefits of data sharing to be sufficient to outweigh their concerns about data sharing.

In relation to ethics, the concerns are:

- That the original ethics approval was based on collecting the data only for the current project
- That re-use by others might enable de-identified human data (for privacy reasons) to be re-identified, e.g. by combining it with other data sets
- For animal research, the publication of metadata related to data collection methods might incite public protest or acts of aggression towards researchers and/or their institution
- For biosafety research, the publication of metadata and data could be misused by terrorists and criminals.

In relation to industry partners, the concerns are:

- The data may be owned by the partner
- The data may contain observations of the partner organisation which the partner would not wish to make available to its customers, competitors and/or regulators.

Researchers usually exhibited a very strong personal sense of ownership of their research data, despite generally being aware that it was legally owned by QUT. Researchers also believed that their research data was part of their competitive “edge”. “Ownership” concerns about data sharing included:

- Whether the original researcher would be given credit for providing the data or whether the recipient researcher might pass it off as their own data, either explicitly or by omission
- The freely giving away of data acquired at considerable cost, including such costs as:
  - The actual costs incurred in collecting the data
  - Years spent building track record and writing grant applications to obtain the funds to collect the data

Sydney, Australia

6th eResearch Australasia Conference 28 Oct-1 Nov 2012
The in-kind researcher time and the institutional overhead not provided through most public grants agencies (that is, it was not all public funding)

- The risk that other researchers could reap benefits from the research data before the original researcher had the opportunity to do so (e.g. publications, grant applications).

Although the objections seemed primarily to focus on making data available to others, concerns were also expressed about receiving data from others, including:

- Data collection is an integral part of research training and therefore research students would not be properly prepared for their careers if they were able to rely solely on the use of others’ research data
- The difficulty of ascertaining that the data was collected in a manner appropriate to the new project and compatible with any ethical requirements of the new project
- The practical difficulty of giving credit to a very large number of original researchers when many data sets have been progressively combined or “mashed-up”.

**Better Motivations for Data Sharing**

If researchers are to be persuaded to share research data, they will require more compelling motivations than the traditional altruistic justifications. We found that, while open access to research data was not attractive to researchers, many responded more positively to the proposal that they provide open access to metadata but negotiated access to the data itself. The benefits of this approach allow a number of the concerns about sharing to be addressed.

For ethics and industry partner concerns, the data owner can offer to run queries over their data for the benefit of other researchers, thus allowing the owner to ensure confidentiality (as required) and/or to ensure that the intended use of the results of those queries would be compatible with the ethical/legal constraints under which the data was collected.

By requiring negotiated access to the data, data owners can convert the concern about competition into an opportunity for collaboration. At a minimum, the data owner could request that access to the data requires the citation of their preferred publication (as citations to publications are generally key performance indicators, whereas citations to data are not). By discussing the research questions of the other researchers, it might be possible to propose a joint paper or joint research project to pursue them. It also allows the data owner to propose that the other researcher share their data, perhaps to form an even larger and more useful dataset. In this way, increasingly large collaborations of researchers could be formed around an ever-growing dataset, providing opportunities for even higher quality research and the opportunity to use those collaborations and shared datasets as the basis of applications for larger research grants and research centres (e.g. Centres of Excellence, Cooperative Research Centres). We found many researchers were very receptive to the idea of advertising their data to promote collaborative opportunities, as the outcomes are generally beneficial to both the researcher and their institution.

If the practice of negotiated access leading to collaborative research activity became the norm, then the concerns about receiving data from others would be mitigated, as the data sharing would occur in the more trusted environment of that collaboration.

**Ethics Processes**

Although researchers were aware that it was possible to vary their ethics approval, the general feeling was that ethics approval was a difficult process and they were unwilling to invest the effort in obtaining a variation to facilitate subsequent data sharing. Realistically, we expect data gathered under an existing ethics approval is unlikely to be shared. Therefore it is important to try to encourage both researchers and the institutions to rethink their approach to future ethics applications to facilitate rather than frustrate future data sharing, a topic which has been addressed by a recent ANDS Guide on Ethics, Consent & Data Sharing [2].

**Recommendations**

As a result of the experiences outlined above, we recommend:

- That the benefits of negotiated access to research data should be explained to researchers, particularly those who appear reluctant to entertain the idea of open access
- That institutions give serious consideration to the ANDS proposals in relation to ethics going forward.

**References**


ABOUT THE AUTHORS

Philippa Broadley is a Data Librarian at the Queensland University of Technology (QUT) Library and is project manager for two Australian National Data Service-funded projects – the RDA Gold Standard Record Exemplars project, which is investigating issues surrounding metadata record quality and connectivity, and the Metadata Store project, which aims to produce an integrated institutional data registry system for storing descriptions of research data produced by QUT researchers. Philippa has previously worked in a number of library positions during her five years at QUT Library, including Liaison Librarian and Library Adviser. Philippa’s interests are in the areas of research data management and open access. Philippa holds a Master of Information Management (Library Studies) from QUT and a Bachelor of Arts (Communication Studies) from the University of the Sunshine Coast.

Kerry Raymond is a Professor of Information Technology in the Faculty of Science and Engineering at the Queensland University of Technology (QUT) and the chair of QUT's eResearch Working Party. Her research interests are in model-driven architecture (MDA) and model-driven engineering. Kerry has contributed to many international standards in the International Organization for Standardization (ISO) in the area of distributed computing and in the Object Management Group (OMG) in the area of MDA. Kerry has previously held roles as Research Leader and Unit Leader in the Cooperative Research Centre for Enterprise Distributed Systems Technology (DSTC) and academic positions at the University of Queensland and the University of Melbourne. Kerry holds a Ph.D. and Bachelor of Science (Honours) both in the field of Computer Science from the University of Queensland.