Navigating new terrain: research data skills in the context of information literacy

CRIG Seminar, 30 November 2011

Sam Searle
Data Management Coordinator, Monash University Library
samantha.searle@monash.edu
Overview

- Characterising research data
- Why research data management is an important issue for universities
- Research data and information literacy
- Research data skills at Monash
  - Getting started: 2009-2011
  - Bridging the gap: 2011
  - Towards a new approach: 2012

- Challenges

- Discussion
Characterising research data (1)

- Original material generated by research and/or existing resources obtained for the purposes of research
- Published resources can be considered as data if they are used as the primary materials for a study, analysis or experiment
- Valuable part of the research
  - Validates the results
  - Enables others to build on findings
Characterising research data (2)

- New (collected or created by the researcher) and/or existing (sourced from somewhere else)
- Qualitative and/or quantitative
- Multiple formats
Common types of research data

- Statistics and measurements
- Results of experiments or simulations
- Observations e.g. fieldwork
- Survey results – print or online
- Interview recordings and transcripts, and coding applied to these
- Images, from cameras and scientific equipment
- Textual source materials and annotations
Common types of research data

- Statistics and measurements
- Results of experiments or simulations
- Observations e.g. fieldwork
- Survey results – print or online
- Interview recordings and transcripts, and coding applied to these
- Images, from cameras and scientific equipment
- Textual source materials and annotations
Part of responsible research

- **Section 2, Managing Data and Primary Materials**

- Sets out standards for researchers and their institutions - data storage, retention and re-use
Funding agencies – international

- US National Science Foundation now mandates a data management plan as part of the grants process
Funding agencies - Australia

e.g. ARC Draft Discovery Projects 2011 Agreement – outlines storage and dissemination expectations for data collected as part of projects

20 Material Produced under this Agreement and Dissemination of Research Outputs

20.1 The Administering Organisation must establish and comply with its own procedures and arrangements for the ownership of all Material produced as a result of any Project funded under this Agreement.

20.2 For any Material produced under this Agreement, the Administering Organisation must ensure that all Specified Personnel (Chief Investigators, Partner Investigators and Fellows):

(a) take reasonable care of, and safely store, any data or specimens or samples collected during, or resulting from, the conduct of their Project;

(b) make arrangements acceptable to the ARC for lodgement with an appropriate museum or archive in Australia of data or specimens or samples collected during, or resulting from, their Project; and

(c) include details of the lodgement or reasons for non-lodgement in the Progress Reports and the Final Report for the Project.

20.3 The Administering Organisation shall consider the benefits of depositing the data and any publications arising from each Project in an appropriate subject and/or institutional repository wherever such a repository is available. If the
Research impact

- Some journals now expect supporting data to be available for peer review and/or publication.
- International trend towards open access to data in archives & repositories (sometimes a requirement of funding agencies).
- Increasing use of multimedia as part of publications and presentations.
- Reaching out to audiences that do not have access to expensive journals.
- Some early evidence that when data is publicly available, the publications get cited more.
Sharing Detailed Research Data Is Associated with Increased Citation Rate

Heather A. Pivovar\textsuperscript{a}, Roger S. Day, Douglas B. Fridsma
Department of Biomedical Informatics, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, United States of America

Abstract \textsuperscript{Top}

Background

Sharing research data provides benefit to the general scientific community, but the benefit is less obvious for the investigator who makes his or her data available.

Principal Findings

We examined the citation history of 85 cancer microarray clinical trial publications with respect to the availability of their data. The 48\% of trials with publicly available microarray data received 85\% of the aggregate citations. Publicly available data was significantly (p = 0.006) associated with a 69\% increase in citations, independently of journal impact factor, date of publication, and author country of origin using linear regression.
Research data and information literacy

- Increasing recognition that improving research data management within universities will require not just policy and technology but also *skills*

- Education and Training Birds of a Feather session is always well-attended at the eResearch Australasia conference, which is where most of the ‘data management people’ gather

- *But… the people at that BoF probably don’t know about you. And… do you know about them?*
Changes in the conduct of research activity to reflect more globally oriented, collaborative and multi-disciplinary approaches necessitate reassessment of the quality of our research training programs.

Priority 4.3: Development of new models for research training that explicitly focus on the professional employment needs of graduates.
Getting started at Monash (2009-2011)

- 2-hour data planning seminar developed by the Data Management Coordinator
- Target: new Higher Degree by Research students
- Branded as part of the Monash Research Graduate School’s exPERT seminar program
- ~150 HDR attendees to date
- Positive feedback, with most finding the class “extremely useful” (26%) or “useful” (60%)
- But lots of areas for improvement
In the meantime… over in Client Services

- Contact librarians and learning skills advisers are partnering with academics
- Work underpinned by theory in the form of the RSD Framework
- Information research skills regularly addressed in unit reviews, skill diagnostics, curriculum design, assessment task design, marking rubrics, etc.
- Emphasis on research skills from undergrad through Honours and on to higher degrees
Bridging the divide: 2011

- Was data management falling down a structural gap between ‘research’ and ‘education’ (university-wide but also reflected in the Library)?
- What if we applied the toolkits and expertise from the research information skills portfolio to the emerging area of research data management?
- Data Management Coordinator joined the IRLS (Information Research & Learning Subcommittee)
- Research Data Management Skills Project will happen under the auspices of this group
From this...

Data Management Coordinator
Information Resources Division

Contact Librarians & Learning Skills Advisors
Faculty teams in Client Services Division

Working with research admins
Research data management stand-alone seminars only

Library

Monash Research Graduate School

Information research skills parallel and embedded programs

Working with academic staff

Faculty

Faculty

Faculty

CRIG Forum: Research Data Skills
30 November 2011 | 17
... to this?
2009-2011
Research data management skills programs have been developed in line with funding agency requirements and guidelines for best practice arising from the Australian National Data Service.

2011—
The new project will continue to be informed by research requirements, but will more closely align research data skills with teaching and learning drivers, including the Australian Qualifications Framework, Researcher Skills Development Framework, and the Monash Graduate Attributes.
2009-2011
Sessions have been delivered face-to-face in classroom settings, with a reliance on ‘talking heads’ and text-heavy slideshows. Although these classes have been evaluated well by students, more varied methods of developing and delivering content are well-established in the Library and would be more effective in building knowledge and skills.

2011—
New programs will take advantage of opportunities to deliver content through new channels, such as the e-learning system Moodle. This would enable the participation of off-campus students and address the needs of researchers with different learning styles. Online tutorials, videos and quizzes could also be used to make face-to-face classes more interactive.
Challenges

- Expanding the concept of information literacy
- Librarian skills
- Scalability and sustainability
Expanding ‘information literacy’

The RIN takes the view that researchers’ knowledge and skills with regards to how they handle, manage and look after their data are an integral part of information literacy - thus ‘information’, in this sense, is taken to include data.

- Research Information Network (UK)
Building skills of new library professionals

Academic and special libraries and some non-traditional employers have an increasing demand for staff who have a good understanding of the research process and evidence-based practice. Australian LIS programs would also benefit from the inclusion of a component on research data management…

- Re-conceptualising and re-positioning Australian LIS education for the 21st century (ALTC project)
## Building on existing staff expertise?

<table>
<thead>
<tr>
<th></th>
<th>Current focus</th>
<th>Extending to RDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find / collect</td>
<td>Search skills</td>
<td>Data capture methods</td>
</tr>
<tr>
<td>Synthesise / analyse</td>
<td>Literature review</td>
<td>Data analysis</td>
</tr>
<tr>
<td>Manage intellectual property</td>
<td>Copyright</td>
<td>Copyright &amp; data</td>
</tr>
<tr>
<td></td>
<td>Fair use of publications</td>
<td>Re-use &amp; licensing</td>
</tr>
<tr>
<td>Organise</td>
<td>Organise published works with tools like Endnote</td>
<td>Structure data and metadata using a range of methods/tools</td>
</tr>
<tr>
<td>Communicate</td>
<td>Writing a thesis</td>
<td>Options for data dissemination</td>
</tr>
<tr>
<td></td>
<td>Publishing options</td>
<td></td>
</tr>
<tr>
<td>Have integrity</td>
<td>Plagiarism</td>
<td>Privacy Confidentiality</td>
</tr>
<tr>
<td></td>
<td>Referencing</td>
<td></td>
</tr>
</tbody>
</table>
JISC Research Data Management Training Materials (RDMTrain)

- Program of projects in which information professionals and academics collaborated on discipline-specific content
  - CAiRO: Curating Artistic Research Output [performing arts]
  - DataTrain [archaeology, social anthropology]
  - DATUM for Health: Research data management training for health studies
  - DMTpsych: Postgraduate training for research data management in the psychological sciences
  - Research Data MANTRA [social science, clinical psychology, geoscience]

- All content is freely available in the project website or in the JorumOpen repository in the UK
Discussion questions

- What are the benefits and risks of expanding the definition of ‘information literacy’ to include skills relating to the creation, use and re-use of data and primary materials?
- What existing strengths do we have that we can build on in this new area?
- How might we collaborate across institutions, e.g. to produce materials in a certain subject area?