

# Not drowning, ingesting: dealing with the research data deluge at an institutional level

Andrew Harrison  
ARROW Librarian  
Monash University Library  
andrew.harrison@lib.monash.edu.au

Sam Searle  
Data Management Coordinator  
Monash University Library  
samantha.searle@lib.monash.edu.au

## **Abstract:**

*Australian government funding and policy guidelines increasingly encourage researchers to deposit their research data in institutional or subject repositories, but there are significant technical and organisational practicalities involved in achieving this. In this paper, Monash University Library staff members with responsibility for repository development and research data management describe their work together to establish the Monash University ARROW Repository as a key part of the university's overall program to improve research data management. Repository-related activities are discussed in the context of wider developments, both in technical infrastructure and in terms of professional development and outreach to researchers.*

## Introduction

Australia is at the forefront of international efforts to promote better management and greater re-use of data created by research. The *Australian Code for Responsible Conduct of Research* (NHMRC & ARC 2007) assigns joint stewardship of data to researchers and their universities. These funding councils also encourage grant-holders to consider the long-term future of their data; from 2010, Australian Research Council Discovery Projects Funding Rules (ARC 2009) and the National Health and Medical Research Council policy on dissemination (NHMRC 2009) ask researchers to consider depositing data and publications in an institutional or subject repository whenever this is possible. In relative terms, the Australian National Data Service (ANDS) – a national programme to improve data practices, infrastructure and services – is unmatched anywhere else in the world.

This policy landscape provides opportunities and challenges for university libraries. At Monash University, data management responsibilities cross unit boundaries and involve the Library alongside the Monash e-Research Centre (MeRC), Information Technology Services (ITS), Archives and Records, the Research Office, and the Monash Research Graduate School (MRGS). The Library has a leadership role and in line with a recent UK report on data management roles and responsibilities (Swan & Brown 2008, 24-25), is strategically repositioning itself in three ways:

1. Expanding the scope of institutional repository services, by
  - a) increasing the quantity and variety of research datasets disseminated via the Monash University ARROW Repository, and
  - b) consolidating the repository's role within a broader technical infrastructure for managing research data at all stages of a project, not just at the point of publication / dissemination;
2. Developing a new professional strand of practice, data librarianship, which builds on librarians' existing expertise and relationships with the research community; and
3. Raising researcher awareness through a *sustainable* programme of outreach and training activities with researchers.

This paper discusses each of these three areas, describing foundations, challenges and issues encountered, and some strategies being investigated and adopted for the future.

# 1a Quantity and types of research data in the Monash University ARROW Repository

## Foundations

Monash University has been exploring the use of the ARROW Repository for research data since 2005. Monash's lead role in collaborative national initiatives like ARROW (Australian Research Repositories to the World), DART (Dataset Acquisition, Accessibility and Annotations e-Research Technologies) and ARCHER (Australian ResearCH Enabling enviRonment) prompted demonstrator projects in disciplinary areas such as X-ray crystallography and ethnomusicology.

## X-ray diffraction images for protein crystallography

Use of the ARROW repository in the context of crystallography has been well documented (Androulakis et al 2008, Clarke et al 2009b). The project has been very successful in allowing Monash University to:

- build relationships between researchers, technical specialists and librarians
- test a model for ingesting large-scale research data into an institutional repository
- highlight repository service aspects like persistent identification / resolution and dissemination via Open Archives Initiative-Protocol for Metadata Harvesting (OAI-PMH)
- suggest ways in which institutional repositories can support discipline-specific search and discovery layers like The Australian Repository for Diffraction Images (TARDIS).

The crystallography images presented a considerable technical challenge to the ARROW repository software, requiring the handling of large gigabyte-scale files with web application software designed for megabyte-size files. In the first instance, a representative object was manually created and loaded by one of the authors. The lessons learned then led to the creation of tools to transform the images and metadata from the lab into repository-ready Metadata Encoding and Transmission Standard (METS) records which were automatically imported into the repository with no staff input. The tools developed by the school and the ARCHER project were very successful in transforming the crystallography data into a repository ready format.

## Kartomi Sumatran collection and Australian Archive of Jewish Music

Collections of music and interviews are exemplars of research datasets that are not generated from lab experiments and measuring instruments, and are more representative of research coming from the Arts and Performance disciplines. There are currently 174 audio recordings from the Kartomi Collection of Traditional Musical Arts in Sumatra and 8 from the Australian Archive of Jewish Music. The Kartomi Collection also includes several videos and a series of slides taken during field trips. These images and video have been converted to downloadable file formats and added to the collection of audio files, along with the field notes.

Providing access to these resources through the ARROW Repository helps the Archive achieve its aims of providing "a focus for research and performance of

Jewish music in Australia, especially folk and cantorial music and encompassing contemporary Australian Jewish music” (Monash University 2009a).

In contrast to crystallography content, these audio collections were added to the ARROW Repository using the default repository tools and the standardised procedures and processes used for adding traditional text-based research publications. The audio content had already been digitised in a previous project by the School of Music and detailed descriptions of the content were undertaken by an archivist in the school.

## **Challenges and issues**

These early efforts have enabled Library staff and researchers to test the boundaries of current repository software and to identify key technical issues that need to be addressed as part of a broader ‘roll-out’ of repository services for research data.

### **Ingest and online delivery of data objects significantly larger in size than common publication output types**

The crystallography collection confronted the ARROW team with its first challenge: managing large digital files for direct download. The repository access portal software is configured for Hypertext Transfer Protocol (HTTP) access to the content via common web browsers and Multipurpose Internet Mail Extensions (MIME) type specific plug-ins. This works well for PDF format articles up to 2MB and stretches for archival quality JPEG2000 files, but has not proven ideal for the crystallography images.

The crystallography collection totals 80 GB for 10 experiments that are stored as several datasets per experiment, with some instrument data files as well. The datasets consist of hundreds of images compressed together as large TAR GZIP (tape archivable files zipped using Gnu utilities) format files. The rationale for this was that the individual images were not in browser-friendly formats accessible to casual searches, and therefore should be packed together as complete experimental sets that researchers could download onto specialised viewing software/equipment available in biochemistry labs. Unfortunately the multi-gigabyte files are not accessible outside the university intranet; staff from other Australian institutions have experienced timeouts when trying to download just one of the sets.

The Library is now starting to break the datasets into individual images for easier access. The aggregating datasets and their parent experiments will be represented as virtual collections in the repository. For users wishing to download the entire dataset there will be a link to TARDIS (The Australian Repositories for Diffraction Images), which offers a browser-friendly File Transfer Protocol (FTP) download of the whole dataset (2-3 GB). The ARROW Repository is not able to accommodate the access functionality developed for TARDIS, which is specifically tailored to one resource type, against the ARROW repository’s requirement to host multi-discipline and multi-format content that is still easily accessible.

## **Accommodation of sets or collections in systems designed for item-level metadata and one-to-one relationships between metadata and objects**

Institutional repository software in Australian institutions has been organised around individual items of research, with the entire contents often categorised as 'eprints' or with separate silo repositories for eprints, digital or theses and image collections. This has presented the challenge of organising and presenting collections in the repository and disseminating them appropriately to many search providers.

Collection management hinges on unique metadata being assigned to the objects in the collection. Collection-level records were not initially planned for the repository but there is now a need to respond to the development of the ANDS Research Data Australia discovery service. ANDS services require descriptions of research content held in the institution in the *Registry Interchange Format - Collections and Services (RIF-CS)*. To contribute to ANDS discovery services, Monash University will provide RIF-CS metadata describing the collection, along with an appropriate canned search that dynamically pulls back individual objects matching the collection criteria.

The principal advantages to these 'virtual collections' are:

- Existing metadata that comes with the collection can be used.
- Hard-coded parent-child links do not have to be added between objects in the repository.
- Although the collection may grow, the collection view is always accurate.
- An object can belong to many different collections. For example, a thesis can be part of local theses collection, a dissemination collection for the Australian Digital Theses service, and a collection of the research outputs for Faculty X, School Y, and so on.
- The collection can be exported easily out of Fedora as METS XML files for migration into other systems.

The chief drawback of virtual collection management is that there must be sufficient metadata to define a search result set that pulls back only those objects that legitimately belong in the collection. That is, the extent that the repository can be organised into many sub-collections depends on the richness of the metadata accompanying each object.

## **The need to augment descriptive metadata schema like Dublin Core with richer discipline-specific metadata**

Capturing richer and domain-specific metadata facilitates the management of virtual collections and the provision of metadata for discipline-specific discovery/search services. The repository software used by ARROW is capable of adding multiple metadata schemas to individual objects stored in the repository: Dublin Core can be used for search and OAI dissemination; Metadata Object Description Schema (MODS) for rich bibliographic description and citation display; and more complex schemas for discipline-specific data, such as the Council for the Central Laboratory of the Research Council (CCLRC) Scientific Data Model XML schema, which Monash used to encode the diffraction instrument settings, the wavelength and image angle data for the crystallography datasets. Any schema, provided that the XML is well formed and has an associated XML schema or document type definition

(DTD) file, can be added to a repository object. Applications like ARROW's VITAL Access Portal that can index metadata from any datastream associated with the object, potentially support displays as rich and complex as can be supported by customisation of the display and the metadata associated with the collections.

Research data is even more in need of rich descriptive metadata, as the dataset will be opaque to non-specialists. Even within a discipline, experts will need a minimum level of information to be able to understand the original context that generated the data. Beyond the minimum level, any metadata that describes the function and purpose of the data is likely to lead to greater discovery and re-use by researchers in other disciplines.

## **Rights management**

Copyright is a complex issue for repositories. Monash University has recently completed a four-year process to put a mandatory theses submission policy in place for most Monash PhD candidates (there are some exemptions, in performance art and music). The process was more protracted than it had to be (certainly from a technical perspective), because of the issues around copyright in theses, particularly for third party materials.

In the case of research data, common publication-related copyright issues may be compounded by other legal and ethical concerns, such as the privacy and confidentiality of research participants, as well as institutional policies around commercialisation.

## **Strategies**

### **Increased focus on smaller and less complex data objects**

Monash University will continue to explore more 'difficult' types of data, such as that worked on with the crystallographers, but this work will be led by the Monash e-Research Centre as it involves technical expertise and resources that are not readily available within the Library, and requires consideration of significant technical changes. An alternative strategy for increasing the amount of research data in the ARROW Repository is to focus on 'low hanging fruit', that is, materials that can be more easily accommodated within the current ARROW Repository architecture and business processes. Part of the rationale for this is that the 'data deluge' is not uniform across disciplines. The report *Investigating Data Management Practices in Australian Universities* (APSR 2008) found that the most common types of data are spreadsheets or databases, followed by documents and reports; many researchers do not generate large quantities of data, with about a third having a total volume less than 1GB; and the most commonly used software for data analysis and manipulation are statistical packages like SPSS and Excel.

While further work is needed to clarify an approach, early discussions with stakeholders have suggested a focus on sets of images, audio and video files with requirements (size, metadata schema) like the Kartomi holdings described above; statistical data in the form of spreadsheets and databases; and research data in the form of documents, such as transcripts and reports of content analysis.

A 'low-hanging fruit' approach would enable the Library to more quickly build up research data deposits across different disciplines, since existing processes and skills will work for much of this type of material and it is possible to draw on the experience of other organisations in archiving this kind of data (e.g. the UK Data Archive and the Australian Social Science Data Archive).

### **Establishing workflows for publicly funded research data**

A further way to increase the amount and variety of data content would involve working with new ARC/NHMRC grant holders to improve data management from the start of a project. As noted above, from 2010 new funding agreements will encourage researchers to consider appropriate repositories as the destination for their data. It will be timely to consider approaching new grant holders with information about the planning and ongoing data management activities that will facilitate the deposit of high quality data in the ARROW Repository at the end of the project.

As new grants are unlikely to increase the amount of data deposited in ARROW (and, by inference, contributed to ANDS' national discovery services) in the short term, a second strand of work may involve gathering information about data generated by past ARC/NHMRC projects that may be able to be included in ARROW.

In both cases, the Library envisages working with existing sources of information about ARC/NHMRC-funded projects; much of this data is publicly available, but further information may be available from the University's research information management system, ResearchMaster. This approach would enable the Library to create more sustainable workflows on a year-to-year basis by building on existing processes developed for publication outputs. It would also enable staff to build up data management experience by focusing on recruiting research data from a more discrete cohort, using well-structured engagement processes as part of a formal project.

### **Build partnerships with subject archives**

Australian research libraries are increasingly participating in the development of subject repositories, such as The Australian Repositories for Diffraction Images (TARDIS), the Australian Social Science Archive (ASSDA), the Aboriginal and Torres Strait Islander Data Archive (ATSIDA), BlueNet - the Australian Marine Science Data Network, the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC), and Australian Agricultural and Natural Resources Online (AANRO). International subject repositories are also often a destination for objects and metadata contained in Australian university repositories; for example, Monash University Library is a partner in the Network of European Economists Online initiative, which will disseminate both publications and data through a subject repository called Economists Online (NEEO 2009).

There is no overarching framework for institutional repositories and subject repositories working together. The Library has identified a need to develop a partnership strategy that would build on the strengths of the ARROW Repository,

whilst recognising that subject repositories can offer significant functionality, particularly around access controls and end user interfaces, that a multi-purpose, multi-formal institutional repository cannot.

# 1b Consolidating the repository as part of a broader research data management architecture

## Foundations

An institutional repository is just one part of the technical infrastructure needed to support better research data management. A diagram developed by Andrew Treloar (*figure 1*) provides a model for the way in which research data may move from private to more public systems, over 'curation boundaries' that can be mediated by researchers and information professionals in different roles.

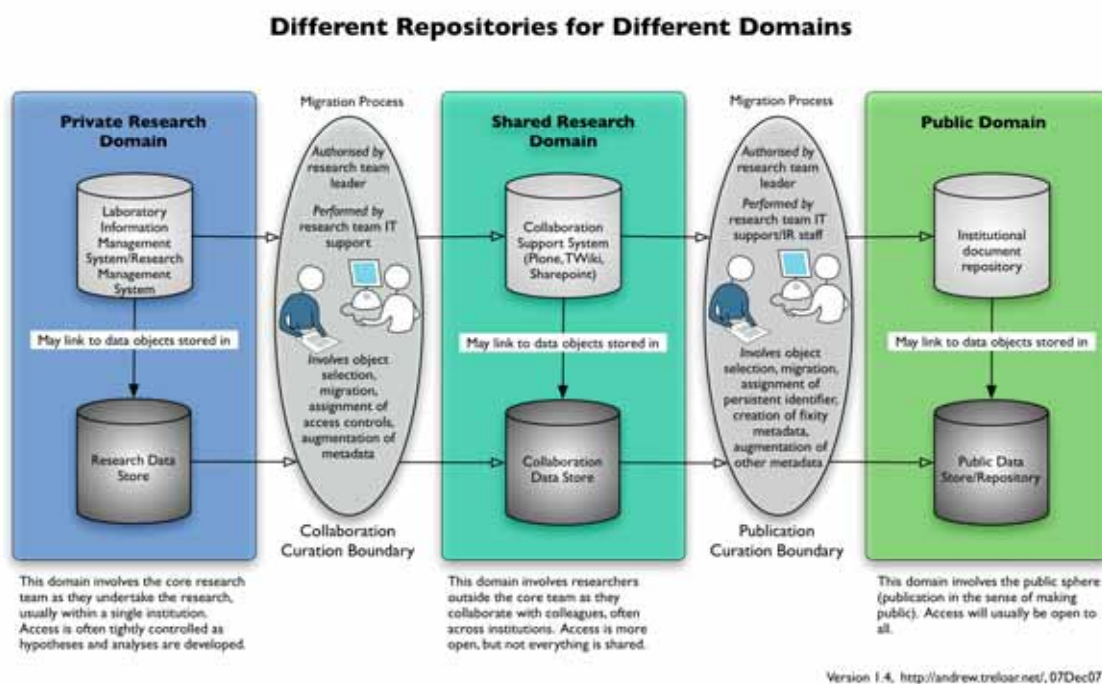


Figure 1

Within Monash University, other parts of this architecture (see *Figure 2*) include the University's Large Research Data Store (LaRDS), collaboration and wiki environments such as Sakai and Confluence, and other tools for managing data and metadata (for example, laboratory information management systems and Mediaflux).

These tools are essential because researchers need to manage data that is not able to be deposited in the ARROW Repository, either because it is not in a final state for publishing or archiving or because access must be temporarily or permanently restricted to meet legal, ethical or commercial requirements.

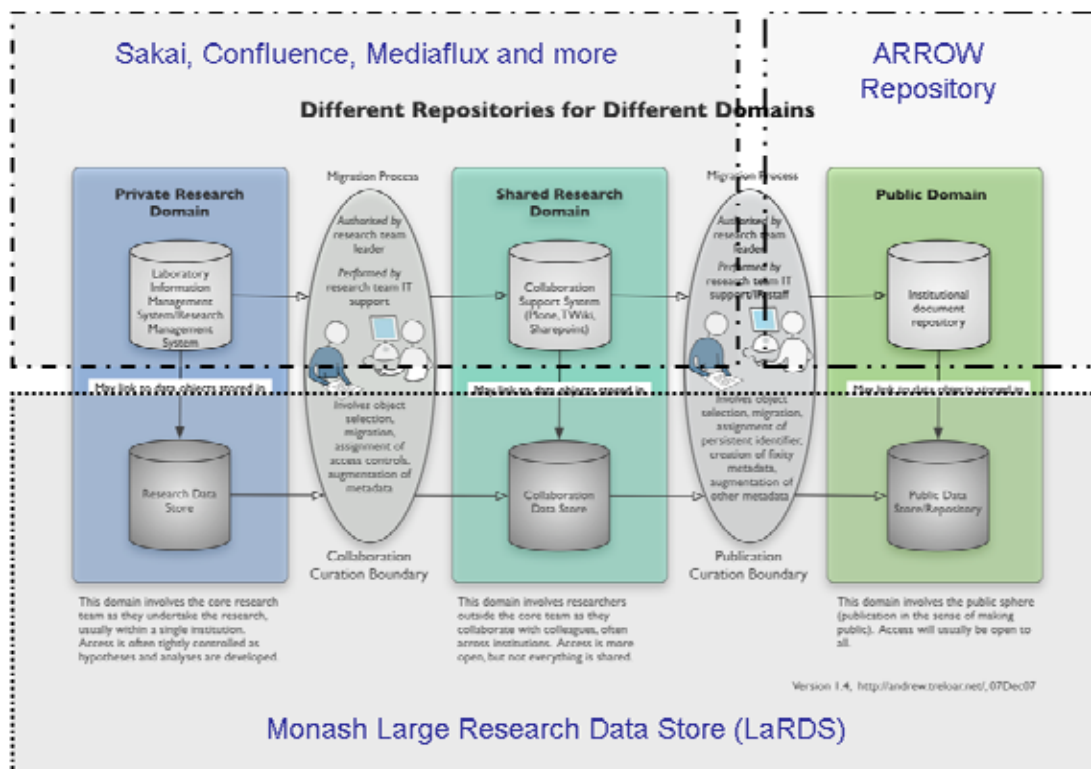


Figure 2

## Challenges and issues

A challenge for the Library is to ensure that the ARROW Repository is well integrated with other parts of the overall technical infrastructure, together with the emerging policies and practices surrounding these. As different parts of the technical architecture have emerged, risks and opportunities associated with purpose, interoperability and rights management have been addressed. Given the ease with which researchers can make data available through other means, it is important to reinforce the strengths of the ARROW Repository as a set of services focused on capturing the University's research output and continually increasing the impact of that research.

## Strategies

The Library is working towards utilising LaRDS to take advantage of regimes of security and back-up appropriate for compliance with the *Code for Responsible Conduct of Research* and with ethical guidelines. Further opportunities are presenting themselves at the researcher's desktop. As tools like Mediaflux – a digital asset management system particularly suited to multimedia research collections – are rolled out within particular research groups, the Library can advise on metadata schema and controlled vocabularies that would facilitate the production of repository-ready content and reduce the effort spent on processing and quality control. A pilot project is under way in the new Department of Architecture, with a 'lite' version of MODS proving to be rich enough to describe architectural research data (images,

models, documents) in ways that make sense to researchers in that field. The immediate priority of the researchers is to use Mediaflux for internal data management, but with support from the Library they should be able to export records that can be more easily ingested as a datastream into the ARROW Repository in future.

## 2 Directions in data librarianship

### Foundations

Early national repository projects like ARROW, DART and ARCHER have offered a new career direction for library staff, many of whom are now involved in ANDS. In 2008, the Library also created a full-time Data Management Coordinator position in the Information Resources Division, where ARROW staff are located.

Monash University recognises that a small number of dedicated roles will not be sufficient to deal with data management in a large and distributed organisation. The Library's first phase of professional development for **a broader group of staff**, through an initiative called Dare, has been well documented elsewhere (Clarke et al 2009a & 2009b). Dare initially provided a forum for a large group of contact librarians to learn about data management through attending voluntarily regular meetings and seminars, but the group has evolved to include Library staff from other areas (such as the ARROW team, Electronic Resources, Learning Skills, and IT support), and collaborators from other Monash University units (e.g. Monash e-Research Centre) and local external organisations (e.g. CSIRO, ANDS).

As a direct result of Dare, Monash University contact librarians increasingly engage directly with researchers to improve the way research data is stored and managed. Librarians consult with staff and new PhD students about their needs and promote data services (particularly planning support) at faculty, department, and research team meetings. Library staff are also encouraged to present at and participate in related events, such as the Victorian eResearch Strategic Initiative (VeRSI) fortnightly eCoffee seminars, the Monash e-Research Centre's annual e-XPO, at which a booth was staffed by volunteers from across the Library, and one-off events such as a seminar by the visiting Data Librarian from the University of Edinburgh (Rice 2009).

### Challenges and issues

Managing the diverse needs of a group of library professionals in relation to a new area like research data management is challenging. Dare has been successful in piquing the interest of a large number of staff; many seek coaching or support to embed particular activities within their existing roles over time, while a smaller number of staff want to expand their data management skills in this area more rapidly or are seeking challenges relating to project management, strategic planning and communications (all important in this area). Because of these diverse needs the Library must now turn its attention to creating additional opportunities for greater levels of engagement by a sub-set of the Dare group.

These internal challenges take place in a broader global and national context in which the future of data librarianship as a professional strand is not clear. An influential UK report has recommended that "the research library community in the UK should work with universities and research institutes to define properly and to formalise the role of data librarians, and to develop a curriculum that ensures a suitable supply of librarians skilled in data handling" (Swan & Brown 2008, 30). In

Australia, the situation is similar; data management does not feature in postgraduate library courses, and career structures are not yet well defined.

## **Strategies**

### **Creating opportunities for team-based project work**

Librarians have started working in project teams with MeRC staff and researchers from areas including climatology and architecture. As part of a multi-disciplinary project team, librarians can provide guidance on data planning, inventories and metadata standards, and contribute to communications and relationship management. MeRC staff focus on technical issues such as storage, data transfer, and the roll-out of software systems for managing and describing digital assets. The researchers drive the process by advising on their requirements and testing systems, as part of an iterative process.

### **Secondment / relief opportunities**

Secondments and relief arrangements are also professional development opportunities. With the support of senior managers, the role of Data Management Coordinator was offered as a Higher Duties Allowance (HDA) opportunity while the incumbent was on leave. Considerable interest was expressed in the temporary position, which offered flexibility to the successful candidate to not only 'keep things ticking over', but to develop a mini-project in an area that combined data management with an existing professional area of responsibility.

### **National capability development**

Building data management expertise is a national issue for research libraries, requiring a coordinated response. ANDS will be a crucial hub for sharing expertise across institutions. Monash University Library intends to share as much knowledge about its processes and practices as possible, through forums including ANDS working groups and discussion boards, and by supporting staff participation in external events with other data management professionals. The Library also plans to expand the content on the research data management website to include more information aimed at colleagues in other research organisations.

### **Working with academic staff to develop a 'special topic' paper**

One reason data management does not currently feature in postgraduate library courses is that the changing nature of the field makes it difficult to establish a curriculum that will meet the needs of individual professionals as well as the research community. Monash University academic staff involved in the delivery of postgraduate degrees in information and knowledge management are investigating the feasibility of a 'special topic' paper, which would provide an issues-based overview for students that are often already engaged in professional positions. The Data Management Coordinator has been involved in these discussions, and has also provided a guest lecture to current postgraduate students to highlight career options in this emerging area.

## **3 Raising awareness through outreach to researchers**

### **Foundations**

Technical challenges sit alongside significant issues relating to people and organisations. Researchers face complex legal requirements, ethical protocols, policy frameworks and technical concerns; their knowledge, decision-making and day-to-day practices and the culture of their disciplines all impact on their ability and willingness to manage research data well, and to deposit research data in a repository. The Library's evolving approach to outreach encompasses many activities.

### **Website**

The highest priority communications activity has been the development of a website aimed at Monash researchers, [www.researchdata.monash.edu](http://www.researchdata.monash.edu). Most of the content consists of a set of guidelines on topics such as storage and back-up, metadata and documentation, retention and disposal, and depositing data in a repository or archive. Templates include a data planning checklist that has been the subject of several phases of testing and feedback with researchers and contact librarians (Searle 2009).

### **Inductions**

Inductions for new research staff and new higher degree by research (HDR) students raise awareness of research data management issues at a time when researchers are potentially more open to new ideas and ways of doing things. There is always a risk of information overload at an early stage in a researcher's time at Monash, so the introduction to data management is very brief and focuses on providing contacts for follow-up and links to further information.

### **Information sessions**

The Data Management Coordinator, often in combination with other staff from the Library or the Monash e-Research Centre, delivers information sessions on request to groups of all kinds. In the past six months, about one session a week has been delivered. Over time, contact librarians will undertake more of these within their faculties.

### **One-on-one data planning consultations**

Some contact librarians have been conducting unstructured interviews to assess data management requirements and provide referrals for storage, management tools and additional expertise. More recently, these consultations have focused on the use of a data planning checklist (Monash University 2009b), which is designed so that researchers can easily identify gaps in their data planning and the potential issues that may arise from this. The checklist provides links to appropriate web-based resources and contact details for support services so that researchers do not hit a 'dead end' if they cannot complete any part of the checklist.

## **Training sessions for higher degree by research students**

Professional development for researchers is a high priority. The Data Management Coordinator runs two-hour workshops with higher degree by research (HDR) students. These sessions expose future research leaders to data management concerns early in their careers and treat data planning as an essential research skill.

Feedback on early sessions suggested a need to provide more hands-on training with some of the tools available at Monash. A second workshop was added, giving an overview of the Sakai all-in-one collaboration environment and the Confluence enterprise wiki, both of which utilise LaRDS. Students learn to create a workspace, upload data and set access controls to their data appropriately.

For more established researchers, professional development opportunities are provided on an ad hoc basis, at the request of the department or faculty.

## **Challenges and issues**

Sustainable outreach strategies will always be a challenge in a large and distributed institution like Monash University. With around 3,600 staff and almost as many HDR students distributed across ten faculties and multiple campuses, there is no one channel that can be used to contact researchers, nor can a one-size-fits-all message be expected to appeal to researchers whose data management issues vary depending on their discipline, career stage and a host of other factors.

Further challenges arise around promoting the benefits of data management to researchers when current rewards systems are so heavily weighted towards publication and do not recognise the work that is involved disseminating data as a research output in its own right.

## **Strategies**

### **Identification of key groups**

Given the size of the Monash University research community, an analysis of stakeholders was needed to inform outreach activities and enable prioritisation. Some of the factors being used by the Library to prioritise engagements with researchers include:

- Career stage, e.g. higher degree by research students and early career researchers may be more open to new ways of doing things; they are also valuable as change agents and will be the research leaders of the future. Staff at the end of their careers also need support to make arrangements that ensure the continuing contribution of their intellectual outputs to their institution and discipline.
- Point of need, e.g. new grant holders, may need help to fully understand and meet data management requirements imposed by funding agencies.
- Alignment with strategic areas identified by Monash University and/or by funding and policy agencies.

## **Working with the willing**

Demand from interested researchers for the Library's services is high, so there is little incentive to spend time trying to convince those who do not perceive a need for assistance or believe in the value of the services on offer. The approach will continue to be to find and work with early adopters or 'champions' and to build a portfolio of real-life case studies where changes in practice have resulted in demonstrable benefits to the researchers.

## **Leveraging existing activities and approaches**

The Library will continue a successful strategy of 'piggybacking' on existing communication channels and programmes of activity. For example:

- The Monash Research Graduate School (MRGS) manages logistics and advertises Library-led seminars for HDR students, which have been branded as part of the MRGS exPERT seminar series, already well-known to students and their supervisors.
- Rather than running separate data management sessions for new staff, the Library is part of the compulsory induction programme run by the Research Office.
- The Research Office has also agreed to include data management information in the information packs routinely sent out to new grant-holders.

## **Data management as part of information literacy and research skills**

In the longer term, the Library may work with faculties to embed research data management within undergraduate coursework and Honours programmes, as has been done with other information literacy and learning skills. The Researcher Skill Development Framework (Willison 2008) offers an overarching framework for starting to think about data management skills within existing models for information literacy. The Data Management Coordinator will explore this further in 2010, with the support of other staff such as the Information Literacy Librarian and the Learning Skills Manager.

## **Conclusion**

Research data management is a major challenge for research institutions. It provides an opportunity for libraries and librarians to take on new roles and responsibilities, capitalising on existing infrastructure and available pockets of staff expertise to deliver new or enhanced services to researchers.

Monash University Library has gained practical expertise over several years and has laid foundations for this work. The Library is now developing strategies that will enable us to move ahead in 2010 and later years with data-related initiatives that will consolidate the role of the ARROW Repository, provide development opportunities for Library staff, and raise awareness of the benefits of good research data management amongst Monash University's diverse community of researchers. The Library will continue to share experiences with professional colleagues for the benefit of the wider library and research sectors.

## References

Androulakis, S et al 2008, 'Federated repositories of X-ray diffraction images', *Acta Crystallographa Section D, Biological Crystallography*, Jul 64 (Pt 7), pp.810-4, viewed 31 August 2009, <<http://arrow.monash.edu.au/hdl/1959.1/43945>>.

ANDS – see Australian National Data Service

ARC – see Australian Research Council

Australian National Data Service (ANDS) 2009, *Registry Interchange Format - Collections and Services (RIF - CS)*, viewed 9 September, 2009, <<http://ands.org.au/resource/techdocs.html#rifcs>>.

Australian Research Council (ARC) 2008, *Discovery Projects Funding Rules for funding commencing in 2010*, ARC, Canberra, viewed 31 August, 2009, <[http://www.arc.gov.au/pdf/DP10\\_FundingRules.pdf](http://www.arc.gov.au/pdf/DP10_FundingRules.pdf)>.

Clarke, S, Harrison, A & Searle, S 2009a, 'Scholarly information repository services at Monash University', paper presented to the International Association of Technical University Libraries, Leuven, Belgium, 1-4 June, viewed 31 August, 2009, <<http://arrow.monash.edu.au/hdl/1959.1/75384>>.

Clarke, S, Macmillan, W, Searle, S, Harboe-Ree, C & Clarke, N 2009b, 'To Dare or Not to Dare', paper presented to the EDUCAUSE Australasia 2009 Conference, Perth, Western Australia, 3-6 May, viewed 31 August, 2009, <<http://arrow.monash.edu.au/hdl/1959.1/76552>>.

Monash University 2009a, *Australian Archive of Jewish Music*, viewed 14 September, 2009, <<http://arts.monash.edu.au/jewish-civilisation/resources/music.php>>.

Monash University 2009b, *Research Data Planning Checklist (Higher Degree by Research Student)*, viewed 14 September, 2009, <<http://www.researchdata.monash.edu.au/resources/datahdrchecklist.doc>>.

National Health and Medical Research Council 2009, *Dissemination of research findings*, viewed 14 September, 2009, <<http://www.nhmrc.gov.au/grants/policy/dissemination.htm>>.

National Health and Medical Research Council & Australian Research Council 2007. *Australian Code for the Responsible Conduct of Research*, NHMRC & ARC, Canberra, viewed 31 August, 2009, <<http://www.nhmrc.gov.au/publications/synopses/r39syn.htm>>.

NEEO - see Network of European Economists Online

Network of European Economists Online 2009, *Economists Online*, viewed 15 September, 2009, <<http://www.neeoproject.eu/>>.

NHMRC – see National Health and Medical Research Council

Rice, R 2009, 'DISC DISC-UK DataShare: data librarians and repository managers working together', presentation at Monash University, 28 January, viewed 10 September, 2009, <<http://versi.edu.au/downloads/Rice-Monash.pdf>>.

Searle, S 2009, 'Research data management: an update on Monash University Library activities', Victorian e-Research Strategic Initiative (VeRSI) eCoffee presentation, 28 May, viewed 31 August, 2009, <<http://versi.edu.au/downloads/SearleeCoffee.pdf>>.

Swan, A & Brown, S 2008, *The Skills, Role and Career Structure of Data Scientists And Curators: An Assessment Of Current Practice And Future Needs. Report to the JISC (July 2008)*, Key Perspectives, Truro, viewed 31 August, 2009, <<http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/dataskillscareersfinalreport.pdf>>.

Willison, J & O'Regan, K 2008, *The Researcher Skill Development Framework*, viewed 14 September, 2009, <<http://www.adelaide.edu.au/clpd/rsd/rsd7>>.

Witt, M, 2009, 'Institutional Repositories and Research Data Curation in a Distributed Environment' *Library Trends*, Volume 57, Issue 2, DOI: 10.1353/lib.0.0029, viewed 12 January 2010, < [http://muse.jhu.edu/journals/library\\_trends/v057/57.2.witt.html](http://muse.jhu.edu/journals/library_trends/v057/57.2.witt.html)>.