2014 Research Publications
Repository Survey Report

28 October 2014
## Contents

- Figures .................................................................................................................................................. 3  
- Introduction ........................................................................................................................................ 4  
- Responses ........................................................................................................................................... 4  
- Key Findings ....................................................................................................................................... 4  
- Detailed Findings ................................................................................................................................. 5  
  - General .............................................................................................................................................. 5  
  - Staffing ............................................................................................................................................. 9  
- Software ............................................................................................................................................. 11  
- Statistics ............................................................................................................................................. 12  
- Workflow ........................................................................................................................................... 17  
- Links to Research Management Systems ......................................................................................... 20  
- Copyright .......................................................................................................................................... 23  
- Data management ................................................................................................................................. 25  
- Services ............................................................................................................................................. 29  
- Promotion/Feedback .............................................................................................................................. 31  
- Funding ............................................................................................................................................... 33  
- Methodology ....................................................................................................................................... 33  
  - Appendix 1: Statistical questions referred to CSAC ........................................................................ 33
Figures

Figure 1: Mandated deposit for all research publications ............................................................... 6
Figure 2: Institutional policies that refer to your RPR? ................................................................. 7
Figure 3: Does your institution have an RPR open access policy? ............................................... 8
Figure 4: Providing RPR services to affiliated Medical Research Institutes .................................. 9
Figure 5: Institutional area/business owner responsible for the RPR ........................................... 9
Figure 6: RPR staffing change since 2013 .................................................................................... 11
Figure 7: RPR software and version do you have currently in use ............................................... 11
Figure 8: Using hosted RPR software ......................................................................................... 12
Figure 9: Repository wide usage statistics .................................................................................. 13
Figure 10: Author level or article level usage statistics .............................................................. 13
Figure 11: Usage statistics displayed by RPR software .............................................................. 14
Figure 12: Deposit statistics ....................................................................................................... 14
Figure 13: Filtering robot hits, RPR staff downloads, etc. ............................................................ 15
Figure 14: Reporting attention metrics from social media tools ............................................... 15
Figure 15: Making use of the statistics ....................................................................................... 16
Figure 16: Deposit of metadata records and full text .................................................................. 17
Figure 17: Deposit agreements ................................................................................................... 18
Figure 18: Researcher identifier schemes .................................................................................... 19
Figure 19: Metadata standards .................................................................................................... 19
Figure 20: Routine checking of publisher author-rights policies ................................................. 20
Figure 21: What research management system is in use at your institution? .............................. 20
Figure 22: Involvement in institutional HERDC and/or ERA data collection processes .......... 21
Figure 23: Integration into the workflow of your research management system? ....................... 21
Figure 24: Direction of dataflow between RPR and RMS ........................................................... 22
Figure 25: Tagging mandated deposited publications .............................................................. 23
Figure 26: Links to Sherpa/ROMEO ........................................................................................... 23
Figure 27: Including/embedding copyright permissions in record metadata .............................. 24
Figure 28: Thesis author permissions ....................................................................................... 25
Figure 29: Data management systems ....................................................................................... 26
Figure 30: Direction of data/information flow between RPR and DMS ....................................... 27
Figure 31: Housing research data metadata .............................................................................. 27
Figure 32: Persistent identifiers .................................................................................................. 28
Figure 33: Providing a data storage service ............................................................................... 29
Figure 34: Financial assistance for Article Publishing Charges ................................................ 30
Figure 35: Discoverability through search engines/portals ...................................................... 32
Introduction

CAUL has conducted an annual survey of research publication repositories (RPR) since 2009. The results of these surveys provide a valuable picture of repository development and management over time and increasingly the information is of interest not only to CAUL members but to other organisations including government departments, and research granting agencies such as the Australian Research Council (see the ARC Open Access Policy). Recognising this interest, and to promote repositories and increase understanding about repository practices and trends within Australia and New Zealand, the 2014 report is open access. CAUL members also have access to the institutional data.

In Australia, research publication repositories are also of interest as a means of assisting with data collections for the Higher Education Research Data Collection (HERDC) and the Excellence in Research for Australia (ERA). Late in 2013, the ARC sought and received the cooperation of CAUL in conducting a survey of all CAUL members about aspects of repository management related to HERDC and ERA data collections. The ARC was invited to comment on the draft 2014 RPR survey and provided responses about their particular areas of interest.

Responses

The 2014 RPR survey was conducted in September closing on 26 September. Responses were received from all eight of the New Zealand libraries and all but one of the Australian libraries, giving a total of 46 responses.

Some comments provided by respondents are included in the report as examples, highlighting the detailed findings. The comments are quoted verbatim.

Key Findings

The addition of new questions to the survey has provided insights into a number of practices not previously identified. The list below provides an indication of these as well as important changes in the results of 2014 compared with 2013.

- Comparable increases were recorded for theses and HERDC publications.
- Nearly two-thirds of institutions have affiliated Medical Research Institutes, and of these, the number of institutions that provide RPR services to the affiliated institutions is nearly the same as those who do not, i.e. just on half.
- DSpace continues to be the most commonly used platform and has retained its popularity over time. Use of Vital and Fedora Primo has declined over time and usage of the other platforms has remained more or less constant.
- About one-third of institutions are unable to exclude from their statistics, robot hits, RPR staff downloads or other uses which do not truly reflect repository usage.
- About one-quarter of institutions report that they are able to report attention metrics from tools such as Plum and Altmetrics.
- Ten institutions report that they are considering using ORCID. At the time of the survey, institutional identifiers are the most commonly used identifiers with smaller numbers also using Thomson Researcher IDs, ORCID IDs or other forms.
As in 2013, Research Master continues to be the most commonly used research system, being used by nearly half the institutions. The big growth, however, is in Symplectic which has almost doubled its number of users in the past year, even if from a small base.

There has been a huge increase in the number of institutions reporting integration of the workflows in the RPR and the data management system, from less than 10% in 2012 to more than 60% in 2014.

Nearly 80% of Australian repositories are involved with ERA and over one-half with HERDC.

More than 60% of repositories reported that they have a facility to tag mandated open access publications deposits and provide an identifier for them.

The need for integration of data management and RPR systems is increasingly recognised.

Over half of the institutions say that they are now providing a data storage service.

About a quarter of the institutions stated that their institution provides financial assistance for publication.

### Detailed Findings

#### General

There has been little change, since 2013, in the number of other types of repositories in institutions. About two-thirds of institutions report that there are additional repositories including those for: research data or metadata, learning materials, images or media, digitised heritage or special collections, research collections, dark ERA repositories or separate thesis repositories.

As in 2013, about 35% of respondents have a long term strategy for their RPR, with a number mentioning the need to take an integrated approach to research data and publications. Some responses mention two year plans or high level plans incorporated in the institution’s strategic plan. Comments include:

- Gather requirements for improvement of the system over the next 2-3 years, including the front-end. This will include reviewing complementary software. Integrate with data management tools. Integrate with HERDC publication management system (Symplectic Elements).
- R&D Office initiated an external review of the Research Administrative System by Ernest & Young in 2014. As the result of the review, a more integrated research management environment will be developed via the ICT Strategy over the 3 years from 2015 to 2017. This will incorporate improved integration between the RPR, the research management system and academic profiles.
- Publications from Research Outputs System (ROS) will be migrated to Works from Q4 2015. Valet (current front end will be decommissioned). No specific strategic document but these changes are documented in ROS project plan.
- We are addressing this now and are in the process of developing a longer term strategy.
- We currently support a range of services including research publication collection, open access support, data collections, theses submission and review and cultural heritage and other digitised materials. The sustainability and effectiveness of our current software solution is being assessed with a view to disaggregating and seeking alternative software.
solutions where suitable opportunities exist. The Library is also actively expanding our digitisation program, support for open access and research data management.

- 2 year plan with a focus on: Integration with research data; Integration with University source of truth systems; Data Citation (DataCite); Research Impact Reporting.
- A review paper was written late in 2012 by Associate Librarian, Information Resources, Systems and Infrastructure. A variety of tasks have been, or in the process of being, undertaken in relation to workflow, visibility and relationship of the repository to the ERA/HERDC collection process.

**Deposit mandates**

There was a marked increase in the number of universities mandating deposit for all types of research publications, with the percentage rising from 16% in 2013 to 37% in 2014. Comparable increases were recorded for theses and HERDC publications. The major Australian research funders are now requiring that publications be made open access wherever possible, so it is not surprising that the universities are following suit with their own policies. Theses, however, are not subject to the same external reporting requirements, so it is interesting that they too are increasingly subject to mandate by institutions.

A new question was added for 2014 covering deposit of ERA publications. While more than half of the Australian universities reported that they are now mandating the deposit of ERA publications, only one of the eight New Zealand libraries reported the same for the New Zealand equivalent, the PBRF.

Comments included:

- The omission of creative works from mandating policies,
- The lack of a mandate as such but inclusion of the concept that only deposited works will be included for research assessment purposes,
- ‘Strong support’ for deposit rather than a mandate.

![Figure 1: Mandated deposit for all research publications](image-url)
**Other institutional policies**

The number of universities reporting other institutional policies that refer to the RPR has increased by approximately 10% since 2013. Policies identified include: open access (most often mentioned), research conduct and publication, intellectual property, research data management, collection development and higher degree theses.

![Figure 2: Institutional policies that refer to your RPR?](image)

**RPR open access policies**

This question was asked for the first time in 2014.

Nearly half of the universities reported having an RPR open access policy. However, it was apparent from the responses that there was some confusion between the institutional open access policy rather than a policy which related particularly to the RPR. If this question is to be asked again in future years, the wording should be made clearer.
Embargo periods

52% of institutions indicated that their RPR policy or guidelines included support for publisher embargo periods. In some cases it is apparent that publisher periods are supported in practice, even though this is may not be stated in the policy or guidelines.

- Although not explicitly stated, the terminology used is “Access through the research repository will be consistent with publisher agreements”.
- We can support embargoes for theses, it is a fully manual process of applying and removing. Future software, under consideration, may support automatically managed embargoes based on system dates.

Thesis open access and student choice

60% of institutions reported that, where thesis deposit is mandatory, open access is also mandatory rather than by student choice although the requirement for open access may not apply to creative works. Comments suggest that open access is not always straightforward and other factors may prevent or restrict open access including supervisor embargoes, third-party copyright content, culturally sensitive content, confidential information or commercial-in-confidence agreements.

Providing RPR services to affiliated Medical Research Institutes

Nearly two-thirds of institutions have affiliated medical research institutes, and of these, the number providing RPR services to them is nearly the same as those who do not, i.e. just on half. The issue of responsibility for ensuring that open access mandates are carried out was not explored in the survey. This relates to where responsibility for grant administration lies, rather than with the RPR itself. Repository staff, therefore, need to be aware of where this responsibility does lie and what responsibilities flow on to the RPR. Several institutions commented that services are offered only to those staff with a direct university affiliation.
Institutional area/business owner responsible for the RPR

Responsibility for the RPR seems to be well-consolidated within the library. In 2009, 87% of institutions reported that the library was responsible and this has grown to 91% in 2014. Over the same period there has been a decline in the category of ‘combination library/research office’ from a peak of 19% in 2010 to 7% in 2014. At no time over the past six years has the research office had sole responsibility.

Staffing

Institutions were asked to provide details of their staffing structure. The results of this question need to be seen in their entirety and have been omitted from this survey report. CAUL members are referred to the institutional data.
Table 1 below shows the number of universities with staff from each area involved in processing deposits and the average number within each of those universities, i.e. 7 universities have an average of .74 acquisitions staff engaged in processing deposits in the RPR. The table shows that most universities have repository staff, an average of over 2, engaged in this task. There is no breakdown available for the category ‘Other’.

<table>
<thead>
<tr>
<th></th>
<th># Universities</th>
<th>Average staffing #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acquisitions staff</td>
<td>7</td>
<td>0.74</td>
</tr>
<tr>
<td>Cataloguing staff</td>
<td>9</td>
<td>0.92</td>
</tr>
<tr>
<td>Liaison/Faculty librarians</td>
<td>3</td>
<td>0.70</td>
</tr>
<tr>
<td>Repository staff</td>
<td>41</td>
<td>2.14</td>
</tr>
<tr>
<td>Lending staff</td>
<td>1</td>
<td>0.05</td>
</tr>
<tr>
<td>Other</td>
<td>7</td>
<td>2.07</td>
</tr>
</tbody>
</table>

Table 1: Library staff responsible for processing deposits

Q15 Has your RPR staffing increased, decreased or stayed the same since 2012?

This is the second year in which institutions have been asked how their staffing has changed. The 2014 RPR survey asked not just if staffing had increased in the past year but whether it had decreased. In 2014, 15% of institutions reported that their RPR staffing levels had decreased. Only 20% reported an increase, compared with 27% in 2013. In most institutions, staffing levels remain similar, as can be seen from Figure 6 below, however, it is to be hoped that decreasing numbers do not become a trend. At a time when funding bodies are increasing their emphasis on open access, staff will be hard-pressed to manage workloads.

Three institutions commented that supplementary staff would be available to help cope with the anticipated workload of ERA in Australia or PBRF in New Zealand.

Two universities commented that staffing has remained the same but duties have changed:

- In comparison to 2013, we have stayed the same even though we lost dedicated staffing that year. This is due to the face that in 2014 we are now working with a "devolved team" model of support and assistance.
- Within the university wide restructure the library was restructured and the position of Repository Manager was removed. Librarian positions became more generalised in focused removing the Liaison focus.
There has been no change since 2013 in the number of institutions reporting that they have undertaken an RPR staffing costing exercise, with only 7% reporting that a costing exercise has been carried out.

**Software**

In 2014, respondents were able to indicate not just what software they use but also whether they have more than one software platform in use. Consequently the figures are not directly comparable with previous years. However, it is apparent that DSpace continues to be the most commonly used platform and has retained its popularity over time. Use of Vital and Fedora Primo declined over time and the other platforms remain more or less constant.

Five institutions, or 11%, changed software during the year, compared with none in 2013 or 2012. Four of these institutions provided further information about the change.
• One university implemented Equella and Kaltura,
• One university migrated from Digitool to DSpace,
• One university is implementing Pure, and
• One university has migrated from Vital to Islandora.

There may be more software changes ahead, with two-fifths of institutions indicating that they are planning to review or change software in the next year. Not all reviews are being done with the intention of changing software as can be seen from the comments below:

• We are reviewing the capabilities of its current software.
• We are investigating new self-submission repository software for non-publication.
• We have commenced a project to provide further integration with Symplectic Elements.
• We are reviewing the current repository requirements of the institution.
• We are reviewing how all digital objects are handled.

In 2013, institutions were asked if they were considering using hosted RPR software, and three-quarters replied yes. In 2014, institutions were also asked if they are now using hosted RPR software. Responses show that one-quarter are now using hosted services and a fifth are considering the issue. Had the same question been asked in 2013, some would have responded that they were already using hosted software: Digital Commons has always been hosted and some EPrints and DSpace users have had a hosted service for some time.

![Figure 8: Using hosted RPR software](image)

**Statistics**
There has been a marked increase in the number of repositories reporting that they display both downloads per month and downloads by country.
The question ‘Does your RPR software display author level or article level usage statistics?’ was asked for the first time in 2014 and reveals that nearly four-fifths of institutions are able to display article level statistics and only about one-third author level statistics. All those institutions able to display article level statistics could display author level statistics. However, there was one institution which can display author level but not article level and nine institutions which can display neither. Making author level statistics available is a useful facility when institutions wish to enable their researchers to identify downloads of their own publications.

Most institutions can display both metadata views and full text documents usage statistics as can be seen from the table below. One-fifth can display full-text downloads only and fewer than that metadata views only. Only a small number of institutions can display the number of referrals to an
open access copy of full-text on an external website and referrals to an open access copy of the full-text on an external website. In an environment where researchers may be depositing their work into discipline-based repositories, an ability to show the extent of this will become increasingly important.

### Figure 11: Usage statistics displayed by RPR software

Statistics about deposit which can distinguish between metadata deposit and full-text deposit are similarly useful. Over four-fifths of institutions report that their RPR can provide deposit statistics for metadata deposit and about three-fifths can report statistics for full-text deposit.

### Figure 12: Deposit statistics

The validity of usage statistics can be confused if it is not possible to exclude robot hits, RPR staff downloads or other uses which do not truly reflect repository usage. In 2014, a question was included about this for the first time, and shows that about one-third of institutions are unable to
exclude such extraneous usage. This may have implications for comparing statistics from institution to institution.

The use of social media is becoming more prominent; as informal means of research assessment, such as *Plum* and *Altmetrics*, are gaining in popularity. Repositories in this environment are expected to adapt to take these into account, and hence a question about this was included for the first time in 2014. At this stage, about one-quarter of institutions report that they are able to report attention metrics from these tools. The pressure to do this is likely to increase, and the trend will be an interesting one to watch. One university reports that it is planning to add this facility as part of impact reporting development. Use of Symplectic will make it easier for those institutions using it.

Usage of RPR statistics seems to be increasing, reflecting a growing recognition of the repository as an institutional asset. From Figure 15 below it can be seen that repository statistics are being...
increasingly used for academic promotion, for reporting to university committees, reporting within the library and other purposes. Institutions do not as a matter of course provide statistics to research funding bodies, but one university reported ‘Don't automatically report to funding organisations but have capacity e.g. they can search RPR by funder or ask for a report.’ Similarly another commented that their statistics may be used in grant proposals.

Q27. How do you use your statistics?

There are some anomalies in this Figure 15; it is difficult to see why only 80% of institutions report that they use their statistics for reporting to CAUL, for example, when all do (or should do).

The range of statistics which the various reporting institutions would like to have is wide, and to some extent reflects their inability to have those already discussed above. Altmetrics tracking was frequently mentioned as desirable, as was the capacity to manipulate the statistics already delivered. The comments below represent some of the other wishes.

- Would like to know the most highly cited author; who has the greatest number of items, etc.
- Processing/turnaround time; number of works by type and OA ratio for each type; number of works by faculty/school/research group with OA ratio; subject area/FOR code breakdown; citation counts; altmetrics
- Number of items that don't have document attached, but have an open access link. Access via those open access links would be good as well.
- Author-level metrics, deposit statistics and trends over time
- Metadata-only deposits. NHMRC/ARC deposits Links out to OA at external sites. Incorporation of citation and alt-metrics
- We are working on the integration of a range benchmarking information at an author and article level using Scopus custom data and Incites services.
- Easily find out percentage of full text (without technical staff having to do special query)
Would like to have the actual numbers of full text rather than having to work it out by calculating backwards from % of full text. Information on records with a link to an open access copy.

**Workflow**

There has been some move towards more self-deposit since 2013, with 63% reporting self-deposit by researchers/authors in 2014 as opposed to 52% in 2013. Deposit by faculty/school administrative staff on behalf of authors has also increased. At the same time there has been a reduction in the number of researchers/authors providing content to repository staff to deposit on their behalf. While deposit from researchers/authors/faculty/school has increased, the level of quality control on the part of repository staff has also increased as can be seen in Figure 16 below.

There is a wide variety of methods used for depositing as noted in the following comments:

- Repository staff will create metadata records for existing publications for new senior academics provided they supply a publication list.
- Material from ArXiv, RePEC etc is gathered manually.
- Deposit of metadata and file of old theses, with author permission.
- Repository harvests metadata based on institutional affiliation from WOS and Scopus and from linked ResearcherID accounts.

<table>
<thead>
<tr>
<th>Q29. How are metadata records and full text deposited in the RPR?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2013</strong></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td><strong>Self-deposit</strong></td>
</tr>
<tr>
<td><strong>Self-deposit by researchers and/or authors - no quality control by repository staff</strong></td>
</tr>
<tr>
<td><strong>Self-deposit by researchers and/or authors - quality control by repository staff</strong></td>
</tr>
<tr>
<td><strong>Deposit by Faculty/School administrative staff on behalf of authors - no quality control by repository staff</strong></td>
</tr>
<tr>
<td><strong>Deposit by Faculty/School administrative staff on behalf of authors - quality control by repository staff</strong></td>
</tr>
<tr>
<td><strong>Researchers and/or authors provide content to repository staff to deposit into the RPR</strong></td>
</tr>
<tr>
<td><strong>Metadata pushed to the repository by software such as Symplectic with or without prompts sent to author to …</strong></td>
</tr>
<tr>
<td><strong>Material is collected by repository staff independently of researcher and/or authors</strong></td>
</tr>
<tr>
<td><strong>Other</strong></td>
</tr>
</tbody>
</table>

Figure 16: Deposit of metadata records and full text

- Majority of the items are deposited into the RMS first and then ingested into the RPR; some items are deposited manually by repository staff. Will soon implement Symplectic which will push both metadata and files into the RPR.
• Pushed from our in-house publication database Minerva.
• Automate sourcing of data from sources such as Web of Science and Scopus.

The number of depositors providing a deposit agreement appears to have declined over the past year. It is possible that more content is being provided through other sources such as WOS and Scopus or via research management systems such as Symplectic which makes this practice redundant. However, there does not seem to be sufficient institutions using Symplectic at the moment to account for the whole difference. Two institutions, which responded no, indicated other ways of obtaining permissions:

• We have no "depositors" but we send a copy of the deposit agreement to authors when we request the full-text from them.
• When we obtain versions from authors for use in the repository we ask permission via email.

![Figure 17: Deposit agreements](image)

The importance of research identifiers is increasing with the number of institutions now using a research identifier scheme. Nevertheless the numbers remain small, with only one-third of institutions utilising IDs of any kind. Institutional identifiers are the most commonly used with smaller numbers using Thomson Researcher IDs, ORCID IDs or other forms. Ten institutions indicated consideration of ORCID.
Prior to 2014, institutions were asked which metadata standard they used, but only given the option of selecting Dublin Core while nominating other standards as ‘other’. In 2014, the question was modified. More institutions nominated Dublin Core, nearly 90% as opposed to 72% in 2013, while substantial numbers are also using MODS, MARCXML and NDLTD. Others include RIF-CS, modified Dublin Core, Digital Commons own standard and the DC application profile.

It seems from the responses provided that more institutions have carried out formal assessments of their RPR workflow since 2013. A number of institutions have carried out assessment in association with significant events such as HERDC, ERA or the introduction of Symplectic.

Increasing emphasis on open access means that repository staff must be careful about observing author-rights policies. More than 90% of institutions report that they are checking policies before making open access publications available.
Links to Research Management Systems

As in 2013, Research Master continues to be the most commonly used research system, being used by nearly half the institutions. The big growth, however, is in Symplectic which has almost doubled the number of users in the past year, even if from a small base. About two-fifths of the institutions use a combination of research management systems, as many as four being in place in any one university. In-house systems are still found in a number of institutions.

Australian institutions were asked if they are involved in the institutional HERDC and/or ERA collections. The answer is most definitely yes, with nearly 80% of repositories involved with ERA and over one-half with HERDC.
One of the reasons why repositories are not involved in either HERDC or ERA may be a lack of integration between the research management system and the repository. There has been a huge increase in the number of institutions reporting integration of the two, from less than 10% in 2012 to more than 60% in 2014. However, there is still some way to go.

Figure 24 below shows that it is more common for data to flow from the research management system to the repository. Use of a single entry point will become more common as both research management and repository systems develop. Users of Symplectic seem to have an advantage here as can be seen in these comments:

- All RPR content comes from Symplectic, except theses and dissertations which are deposited directly into the RPR by students.
Researchers or their proxies will deposit metadata or metadata and full-text into Symplectic Elements to meet the various reporting or public access requirements for HERDC, ERA and NHMRC/ARC. Where necessary, metadata or metadata and full-text will be pushed through to the Repository.

Figure 24: Direction of dataflow between RPR and RMS

Research funders want to ensure that mandated open access publications are made available, and having a facility to tag these deposits and provide an identifier for each, would be a significant asset. In 2014, more than 60% of repositories had this ability, a small increase on 2013. Some responses indicated that this has to be done manually. Other comments included:

- We create and provide a CAUL/ANDS recommended purl that resolves to the RDA record.
- Data is captured but there is no designated field for this. It is captured in a text field.
- Field available for this information in RPR if grant identifier is known/provided.

Two institutions provided links to examples:

- [https://minerva-access.unimelb.edu.au/handle/11343/41413](https://minerva-access.unimelb.edu.au/handle/11343/41413)
Copyright

The two charts below show that more institutions provide links to Sherpa/ROMEO through their RPR than through their RMS. The charts may, however, be misleading, as some ticked ‘unsure’ or did not reply to the second part of the question. The importance of linking to SHERPA/ROMEO has declined with funder open deposit mandates. More useful would be, SHERPA/FACT (Funders’ & Authors’ Compliance Tool) which is a tool to help researchers check whether or not the journals in which they wish to publish their results comply with funders’ requirements for open access to research. This does not incorporate Australian funders at present, but could be adapted to do so.

While nearly all institutions store information about copyright permissions, there is a huge variety of ways of doing so. These include:

- Wordings specified by publishers are included in the Note field
- Only for NHMRC/ARC mandated material
- Rights statement included with the publication record
- Embedding in a pdf cover sheet
Specific permissions may be referred to in private comments in IR but not the documentation itself.

Always includes a copyright statement, which includes statements required by publishers/copyright owner. Will include Creative Commons licence when applicable. We may include private internal information in private notes.

We include copyright permissions information in the comments fields of the Library Open Repository. In the ORS system we include detailed permissions information and links to publisher policies.

More institutions in 2014 reported that they include/embed copyright permissions in the record metadata, an increase of more than 10%. Not all of this information is publicly available.

Figure 27: Including/embedding copyright permissions in record metadata

One-half of 2014 respondents indicated that they had received between one to five take-down notices/requests in the past year. Of these, only 5 notices, or 11%, had come from publishers. Most requests to remove items were not associated with copyright infringements or legal action, but involved researchers/students wanting to publish books, protect sensitive information, avoid ongoing plagiarism or having other concerns.

Nearly all repository staff (95%) respond to copyright questions, and more than 85% say that the university copyright officer does. Few research office staff are involved, less than 10%. In 2014, the question included a section for liaison librarians and 20% responded yes to this.

Responses to this question show no change between 2013 and 2014, with nearly four-fifths of institutions indicating that they add or have added theses retrospectively to their repository. One way of doing this is to use interlibrary-lending requests to digitise theses and then to add them to the RPR. A number of repositories indicated that they have carried out retrospective thesis projects in the past and these are now complete.

- 2010-2013 we carried out the retrospective submission of all doctoral theses  2014 - working on master theses
1. Library's Digitization Service processes ILL requests for print theses under Copyright Act. Library's Digitization Service is undertaking larger scale digitization of second copies of theses from specific disciplines. Metadata and digital object [restricted] are then entered manually into the IR by IR staff. About to embark on bulk ingest.

2. Approximately 600 Higher Degree and Research theses. Digitised by an external company then added to RPR. Were removed if requested.

Fewer than half of the institutions seek permission in a systematic way. The question does not ask whether the theses are made available on open access, and such a question would possibly elicit a different response. The response from one university makes this clear: ‘By default, digitized theses are metadata records with the digital file restricted [unavailable for access other than to IR staff]. Permission to make OA is sought from the author.’

![Thesis author permissions](image)

**Figure 28: Thesis author permissions**

**Data management**

RedBox continues to be the most popular of all the data management systems in use in the various institutions, with VIV/VITRO next. Despite chatter about FigShare, it has only at this stage been taken up by one institution, although is under evaluation by others. The number of institutions saying that they have no data management system in place has declined, so that now only about one-third are lacking one. Among the others named are Fedora, MINT (for name authority control) and DSpace.
While institutions may be introducing data management systems, fewer than 20% are integrated with the RPR, little change from 2013.

Those institutions where the data management system is integrated, four report that data flows from DMS to RPR and four from RPR to DMS.

- ReDBox metadata records describing research datasets get harvested by DRO.
- Data flows both ways but no single entry point. Link in RPR to DMS if there is a data record. DMS has a link to associated publication records in the RPR.
- A data management project is underway and it is likely to be a single point of entry which feeds into both.
- All research data collections/datasets are pushed from Redbox to ResearchOnline (RPR) and harvest to RDA is done from ResearchOnline.
- VIVO records for datasets include URLs for related publications in the RPR. RPR can include link to relevant VIVO record. Links are added manually.
- Our implementation model of RedBox is to capture and store the collection metadata and then publish the metadata record from RedBox to the VITAT/Fedora institutional repository. When the record is published to VITAL/Fedora the RIF-CS, MARC and DC datastreams as well as the ReDBox form data are pushed to the repository.
The notion that the RPR can be used to store research data metadata is gaining in currency, with more than one-third now doing this. More are planning to do so and it is apparent that ANDS has had a role to play in encouraging this. The number not planning to store research data metadata has not changed over the past three years and remains steady at about 40%.

- A few examples of research data have been collected in previous year’s research data management projects (ANDS funded). Current repository software doesn’t support efficient deposit of research data or dissemination of research data metadata. Monash is looking at future repository software to address this gap in capability.
- A small but growing series of research data sets are described, as part of ANDS funded projects.
• All our datasets records are currently in the repository and in RDA (ANDS)
• Data management is in early stages of development, the University Library has indicated that the RPR can be used to provide access to published datasets, although none have yet been included.

Most institutions are using handles and about two-fifths are using DOIs. DOI use would seem to have declined since 2011. There seems to be some confusion in responses to the question, which is in a section called Data Management, as to whether the DOIs referred to are for data or for publications. Some comments are specific that the DOIs referred to are for publications; others not. Two universities plan to mint DOIs in the near future.

![Figure 32: Persistent identifiers](image)

More than half of the institutions say that they are now providing a data storage service. In some cases this takes the form of on-campus storage and in other cases, storage is outsourced to organisations such as Intersect. Institutions seem to have adopted a variety of approaches to data storage as can be seen these comments:

• The university has a Research Data Store. Only a handful of research data is stored on the RPR.
• Referral to RDSI storage and provision of interim storage with a view to a data repository
• 1. Self-Service private cloud storage: [https://research-storage.griffith.edu.au/](https://research-storage.griffith.edu.au/) (available from 1 August). 2. Archive storage - by consultation. 3. Storage associated with scientific instruments and high performance computing - by consultation. 4. Other specialised storage services eg. database storage, targeted backups - by consultation. 5. Self-managed network drive shares for research projects/groups (currently in pilot, should be in production by end 2014)
• University's IT provides various storage solutions for research data
• Research Data Online which is a DSpace repository, this is harvested into VIVO and then Research Data Australia
Services
The number of institutions providing personal services has changed little since 2013 and remains at about 40%. The most popular services involve the generation of information likely to be of direct value to authors as can be seen by these comments:

- Automatically generated pages for authors (unique ID), Colleges/Centres and Research Codes (FoR and SEO). Can produce publication lists for each of these. RPR is used to populate the publications tab in Research Portfolio for author / researcher page.
- These are automated publications lists generated from the research publications management system (Symplectic) not the RPR to researcher profiles on the website (Research Portfolio Manager).
- Author home pages are generated in both the repository and using repository information on the University website. Authors can also generate code (html, XML, json, rss) to embed in other websites.

A wide range of future services are in planning. These include using the repository as a host for e-publishing initiatives (with or without using OJS), improvements in statistics available to authors and linkages of publications to author profiles. Examples include:

- Automatically generated publication lists for staff pages
- Request a copy button, Almetrics badge, Scopus citation badge
- Researcher self-submission of their data to future large scale storage solution. Easy DOI and ORCID identifier creation.
- Browse by author page which collates all the author’s publications.
- Researcher profiles automatically generated with publications feed from Swinburne Research Bank. This service managed in-house.
- Author ID management (ORCID, ResearcherID and Scopus IDs) and more performance metrics
- Request a copy; Data Curation Health Check; Sherpa lookup.
There has been some controversy around whether or not researchers should have to pay for publication, and most recently, whether they should pay in association with a guarantee of open access. The question did not distinguish between the two.

About a quarter of the institutions stated that their institution provides financial assistance for publication. The comments show that the source of that funding for payment varies, being sometimes the library and sometimes elsewhere. Some institutions are developing policies about article processing charges (APC) and defining where costs should come from. Others are possibly engaged in practices only known to local areas. The comments below reveal that not all repository staff are aware of what is going on, and may provide links to policies or guidelines

- Gold Open access only
- But not by the Library - Faculties pay.
- Some Schools provide financial assistance
- The Library does not provide financial assistance, but there may be some financial support by Faculty/Divisions for APC. The Open Access Policy encourages staff to use available grant funding to publish in Open Access publications including payment of Article Processing Charges for Gold Open Access publishing media (e.g. journals) and for HERDC and ERA eligible publications.
- BMC membership
- Discounted APCs because of Big Deal package subscription, supporter memberships etc.
- Research Office provides funds which the Library manages on their behalf. Is only used for Biomed Central.
- The Office for Research funds Biomed Central and Wiley Open Access APCs. The library administers the payments and submissions.

**Q55. Does your institution provide financial assistance for payment of Article Processing Charges (APCs)?**

<table>
<thead>
<tr>
<th></th>
<th>0%</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>60%</th>
<th>70%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100%</td>
</tr>
<tr>
<td>Unsure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 34: Financial assistance for Article Publishing Charges*
Promotion/Feedback

About one third of institutions report using tools for feedback, much the same as 2013. Feedback takes the form of surveys or email feedback.

Feedback may be used to refine and improve functionality, layout, workflows and deposit processes, prompt the updating of software, assist with future planning, and, most importantly, influence decision makers.

There does not seem to be a repository in the country which does not identify challenges. Some of these are technical and relatively easily addressed. Others relate to the role of the RPR and the lack of recognition or priority it is given in the wider academic community. A lack of understanding of the benefits of open access and whether this is reflected in local mandates is of concern to some. Many of these are long-standing issues.

Comments include:

- Multiple points of entry create frustration for users and is difficult to promote. Apathy - users happy for repository staff to deposit on their behalf, but they do not want to be involved in the process.
- Addressing the concern among many academics that it all involves more work.
- Apathy. Genuine suspicion that we would somehow be breaking copyright.
- Business drivers are around HERDC and ERA, and issues around Open Access get lost in meeting these initiatives within restrained budgets.
- Publications are collected through the research office on behalf of academics so library promotion has focused on identifying open access potential collections of open access research data, images and audio. Promotion has been limited to date due to the manual workflow process being quickly saturated by a handful of interested researchers and library collection specialists. Promotion of open access publications is complicated because it is a faculty and research office activity and there is some resistance to the perceived additional work involved.
- Having enough staff available. Our staffing profile has been a moveable feast for the past few years and it is hoped that it will become more stable over the next 12 month so that the Coordinator can seriously tackle marketing of the repository.
- Lack of service visibility. Highly complex, manual & slow resolution of OA, Google & Google Scholar indexing is flaky & unreliable. Old/out of date software limits our ability to meet expectations and requirements of our stakeholders and users. Limited resources for staffing and technological upgrade/change. Breaking down knowledge/expertise silos between Library Research Services and University Research Office for collaboration.
- Lack of knowledge / interest in copyright and other Intellectual property
- The challenges we have faced are those typical of those facing IR's in large institutions where a] submission to IR is voluntary and b] govt [HERDC] reporting requirements are handled by a separate, siloed system. From 2015, this picture may change as siloing will cease and we will be able to present a coherent message to researchers across a broad range of reporting and deposit requirements and opportunities.
- Acquiring post-print copies or articles remains an issue, with academic staff having a limited understanding of what a post-print is and why we require it.
• Not a lot now. Still some resistance to self-deposit, but benefits, especially for ERA are now well understood. A good, and robust statistics rewrite this year has helped a lot!

The level of discovery through various search engines and portals has changed little since 2013. Almost all RPRs are searchable though Google and over three-quarters via Trove, Google Scholar and ROAR. Library resource discovery services (such as EDS) were included in this question for the first time and about three-quarters of institutions are searchable in this way.

Having the RPR discoverable by a variety of search engines and portals increases the likelihood that contents will be found. However, not all search engines and portals are equal in value. Google, for example, is almost universally used within the user community, so it is reasonable to assume that the most effort should be put into ensuring that RPR contents are available there. Services such as ROAR are seldom used, so the same priority may not be given to it. Google Scholar sits somewhere between. Trove represents an interesting contrast as its content is Australian, so inclusion in Trove adds to the discoverability of Australian content. Trove has an interest in covering all Australian research content and is keen to ensure that it can identify the content of all Australian repositories, but Figure 35 below suggests that it has some way to go before achieving that.

Having content available through the library resource discovery service is an added benefit for each institution. For those institutions indicating that RPR content is not discoverable in this way may reflect a configuration issue or a decision not to include RPR content. The following two comments are relevant in this regard:

• Could also be harvested by EDS but this is not currently activated, so that results are not flooded by ACQUIRE submissions.
• Primo can index the repository but we choose not to as not all content is available and it would produce multiple records

| Q59. Is your RPR discoverable with any of the following search engines/portals? |
|-------------------------------|---|---|---|---|---|---|---|
| Google                        | 80% | 80% |
| Trove                         | 80% | 80% |
| nzresearch.org.nz             | 40% | 40% |
| Google Scholar                | 60% | 60% |
| OAI Search                    | 60% | 60% |
| OAIster                       | 60% | 60% |
| ROAR                          | 60% | 60% |
| Library resource discovery service e.g. EDS | 20% | 20% |

Figure 35: Discoverability through search engines/portals
Funding
More than 90% of RPR are funded through the library budget, which represents no change from 2013. Few receive funding from other sources, only about one in ten. Additional funding has been received to support ERA or other specific projects, from ANDS for data-related development or as support for the implementation of Symplectic repository tools. One university reported receiving funds from the Office of Research Services which is transferred to the Library for the administration of the repository.

Methodology
The Research Publications Repositories (RPR) survey was extensively reviewed in 2014 by a panel including Fides Lawton (ACU), Martin Borchert (QUT) and Steve Cramond (UMelb) taking into account comments following the 2013 survey. The resulting survey is longer, with 61 questions. While some questions were added, two statistical questions (with multiple parts) were referred to CAUL Statistics Advisory Committee for inclusion in the annual CAUL statistical survey (see Appendix 1 for details). Care was taken to ensure that wherever possible comparability of survey data from year to year would be maintained.

Appendix 1: Statistical questions referred to CSAC
Number of items available from the RPR? (Do not include items waiting to be processed)

Number of items available in full-text from your RPR:

- Open access: publicly available
- Open access with embargo: publicly available after a certain period
- Institution Access: available for users within our institution
- Open access: archived but not available at all
- Peer reviewed published research outputs available on open access
- Peer reviewed published research outputs not available on open access
- Peer reviewed published research outputs available for users within your institution only
- Non-peer reviewed published research outputs available on open access
- Non-peer reviewed published research outputs not available on open access
This report was prepared by Margaret Henty, CAUL Communication and Policy Officer, on behalf of the CAUL Research Advisory Committee and the CAUL Executive.

Contact the CAUL Office at caul@caul.edu.au or +61 2 6125 2990

www.caul.edu.au