CAIRSS 2011 Repository Managers Survey

Introduction

The CAIRSS 2011 Repository Managers Survey was held in May 2011. The response rate for the survey was 82.2% (37 out of a possible 45 respondents). In 2011, the Survey also included the new CAIRSS members – 5 CONZUL institutions. The survey questions focused on several areas including general repository questions; repository staffing; repository software; repository statistics; repository copyright; mandates and policies; research workflows; repository perceptions; data management; and CAIRSS feedback.

Questions for the 2011 Survey were based on the 2009 and 2010 Survey. This was done so that trends may be established in some of the areas. New questions were added where new areas of interest had emerged in the approximately twelve months between the 2010 Survey and the 2011 Survey and included a section on “non-CAIRSS collaboration”.

Methodology

The 2011 Survey was conducted as an online survey. Prior to conducting the Survey, approval was sought from the USQ Human Ethics Committee to collect the survey data. The application submitted stated that the data collected would not be released to the public in any identifiable manner, and that any data released would only be done so once all identifying information had been removed / replaced.

The CAIRSS Advisory Committee was given the opportunity to provide feedback on the Survey questions, and feedback received was integrated into the Survey structure. This included providing a new section on “non-CAIRSS collaborations”, as well as the inclusion of new questions in other existing areas.

Once approval had been given by the Ethics committee, and the online survey had been constructed with the incorporated feedback from the Advisory Committee, all CAIRSS repository managers were invited to participate on a voluntary basis. The invitation also asked each participant to indicate that they agreed to participate in the Survey, to fulfill requirements of the ethics approval.

Participants were given several weeks to complete the survey, and also invited to post any queries or requests for clarification about the Survey to the CAIRSS Repository Program Coordinator. Reminders were sent to all participants at regular intervals during the survey period.
Findings

The response rate from CAIRSS members in 2011 remained roughly the same as 2010 (82.2% in 2011, 82% in 2010), however the universities who chose not to respond were different to the non-respondents in 2010. This has somewhat skewed some of the figures, and this will be mentioned below.

The response rate from the new CONZUL institutions was 100% - a fantastic result. These institutions also reported that they were quite pleased to be able to participate in the survey this year.

General

Number of records in research repository

A steady increase in the amount of content in repositories can be seen. While the majority of research repositories still reported between 5,000 and 9,999 records (41%), Figure 1 below shows that the figure has increased slightly from 2010 (38%) and that the percentage of repositories reporting numbers higher than 10,000 has increased – 44% of all respondents reported numbers greater than 10,000, up from 28% in 2010.

Figure 1 Number of records in research repository

<table>
<thead>
<tr>
<th>Year</th>
<th>0 – 999</th>
<th>1,000 – 4,999</th>
<th>5,000 – 9,999</th>
<th>10,000 – 19,999</th>
<th>20,000 – 29,999</th>
<th>30,000 +</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>15%</td>
<td>41%</td>
<td>22%</td>
<td>9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>25%</td>
<td>38%</td>
<td>6%</td>
<td>9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>46%</td>
<td>17%</td>
<td>14%</td>
<td>3%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Percentage of Open Access records

A point of interest for the CAIRSS Community is the climbing rate of Open Access records in research repositories in 2011 compared to 2010. When the CONZUL institutions are removed from the responses, Figure 2 shows that in 2011 the percentage across Australian universities has increased to 37%, up from 33% in 2010. This may be attributed to the completion of ERA 2010 and the shift in focus back to recruiting Open Access content to the research repositories. The upcoming ERA 2012 may however impact upon this figure again, something which CAIRSS will continue to monitor. If the CONZUL institutions are included, the figure of Open Access content jumps to 44% across all CAIRSS members. Across CONZUL institutions alone, this figure sits at 91% Open Access content.

Figure 2 Percentage of Open Access Records in Research Repository
Long term preservation strategies

Several questions were added to the 2011 Survey around long-term preservation strategies of institutional repositories. The figures (Figure 3) reported of institutions who have no long-term strategy were somewhat surprising – 50% reported no visible long-term preservation strategy, with 27.8% unsure as to whether one existed, and only 22.2% reporting that they do have a strategy. This is an area which respondents also highlighted as being important, so it is recommended that this is considered as an area of focus for CAIRSS during the next 18 months.

Figure 3 Does your Institution have a long-term preservation strategy?

Staffing

Respondents were asked which area within the institution the research repository sits. These figures (Figure 4) have remained fairly consistent since 2009, although it is possible to see the emerging involvement of the “Office of Research” in conjunction with the Library.

Figure 4 University areas responsible for research repositories
Respondents were also asked as to whether their institution had undertaken research repository staff costing exercises, as this is an area which has been discussed several times on the CAIRSS community eList. Surprisingly perhaps, the majority of institutions (63.9%) reported no specific repository exercises had been undertaken (Figure 5).

**Figure 5 Has your institution undertaken any research repository staff costing exercises?**

**Software**

In looking at where the varying participants may have affected the outcomes of this survey, it is particularly worth noting the question relating to who uses which software platform. The figures generated from this question must be used with care: in particular, it is worth highlighting that figures reported for use of the VTLS “VITAL” software in 2010 indicated 32% of responding institutions to be on that platform. However, in 2011 this figure drops to just 14%. However, when the actual respondents are correlated with the 2010 respondents, it can be found that this number has dropped because of the difference in non-responding universities between the two years. If the same universities had responded both years, it is anticipated that the figure would have remained much the same.

**Figure 6 Software in use within institution**
*NOTE: a) Voluntary respondents between 2010 and 2011 differed, so figures representing certain software increases / decreases do not reflect this, and b) 2011 figures include New Zealand CAIRSS members.

Statistics

Survey respondents were asked to indicate what types of statistics are currently in use with the research repository. Results show that while quite a number of institutions have some sort of end-user statistics in place, fewer institutions make use of management or back-end statistics. Repository statistics has been an area of interest for CAIRSS, resulting in content being created for the CAIRSS website around the area. As this is a dynamic area however, it is something which CAIRSS should continue to monitor.

Figure 7 Are End-User and / or Management/back-end statistics in use in your research repository?

Links with Research Systems

With various methods of reporting now impacting upon research repositories, respondents were asked whether their research repository is actually integrated into the workflow with the institutional research management system. Two-thirds (64%) of the respondents indicated that the research repository was integrated into the workflow.

Figure 8 Is your institutional research repository integrated into the workflow with your research management system?
Further, respondents were asked to indicate in which direction the data flows between the research repository and the research management system. Figure 9 indicates that currently most institutions move data from the research management system to the institutional research repository.

Figure 9 Direction of flow between research repository (IR) and research management system (RMS)

Copyright

One of the questions asked in each year’s survey is whether respondents use OAKList. With 2011’s results, this figure has demonstrated a steady growth through the three years: from 82% of respondents in 2009, through 87.5% in 2010, to 91.7% in 2011 (Figure 10). This increase demonstrates the ongoing value of the OAKList to the institutional research repository community. At the time of writing, the future of the OAKList is not certain, but these figures demonstrate that its value to the community is increasing and that future directions should take this into account.

Figure 10 Do you use OAKList?
Respondents are also asked to indicate the existence of deposit mandates that exist in the institution. Comparing the following two tables (Figure 11 and 12), it is possible to see the imbalance between mandates around theses versus “all research”. The large majority of institutions reported that while they do mandate deposit for all theses, this is not the case for all research.

**Figure 11 Mandated Deposit for all research?**

**Figure 12 Mandated Deposit for theses?**
Data management

As data management continues to emerge as an area of interest for institutional research repositories, the survey questions have been extended. Respondents were asked as to whether they are involved in data management activities at their institution. The 2011 figures (Figure 13) indicate that most research repositories are not involved in data management activities. A number of the negative responses indicated however that this was still under consideration as their institution had not yet decided on a strategy for how to approach data management.

Figure 13 Is your research repository involved in data management activities?

Most respondents indicated further that the research repository would not be utilized for housing research data metadata, and several respondents indicated that investigations were still underway. Some respondents indicated that there would be an anticipated linkage between whatever system houses the research data, and the research repository, but as yet the scope of this is unknown.

As data management issues continue to emerge and it takes on more prominence, it is an area that CAIRSS should maintain a watching brief on, and where possible disseminate information to the CAIRSS community.

The survey asked respondents to indicate what (if any) persistent identifiers were in use in the research repository. Respondents were permitted to select more than one option, as different types of persistent identifiers are often found in use for different purposes. Figure 14 indicates that the percentage of institutions utilizing some sort of persistent identifier has continued to increase from 2009 to 2011. In accompanying comments, a number of respondents indicated that DOIs were in use solely because of publisher requirements, and that it was a definite practice to include them as part of article citations.
Handles on the other hand, were not commented to be used for the same reasons, but rather as something which the institution had chosen to implement.

**Figure 14** Have you implemented any of the following types of persistent identifiers in your research repository?

---

**Institutional research repository perceptions**

Survey participants were asked to comment on the perception of the institutional research repository. Comments received were largely positive, with many indicating that the ERA 2010 had raised the profile of the repository within the institution. A growing acceptance and enthusiasm about the repository was also mentioned, and most positive responses also indicated the value of a ‘champion’ within the academic community who supported the repository. The perception of the repository as a tool for reporting also seems to have emerged, with comments indicating that many academics support and accept the repository and its role in HERDC reporting.

A number of respondents indicated that they still feel there is a large amount of work to be done on promoting the repository. Some of these respondents reported that this was because they are still in the early stages of establishing the repository within the institution, while others commented that they struggled with a lack of support from within the institution or a variance in priorities between areas responsible for the repository.

Respondents were also asked to indicate whether they felt that the perception of the institutional research repository had been influenced by the ERA, by HERDC, or by ANDS-related activities. Respondents were permitted to select as many as appropriate. Figure 15 indicates that the majority of respondents felt that the ERA 2010 activity had influenced the perception of the institutional research, while almost 40% felt that HERDC had influenced the perception, but only 13% felt that ANDS-related activity had had any impact upon the perception of the repository within the institution.

**Figure 15** Institutional perception influenced by ERA / HERDC / ANDS? (Yes)
Non-CAIRSS Collaboration, Support and Linkages

The newly added section regarding non-CAIRSS collaborations provided some points of interest which will be explored further: Most respondents reported that they are members of a repository software group (71.4%). Out of those respondents, 57.7% indicated that they find membership in these groups “Helpful”, as opposed to “Very Helpful” (23.1%) or “Not Very Helpful” (3.8%).

The majority of respondents following between one and three email lists and RSS feeds (67.6% and 47.6% respectively).

Interesting to note however, is that almost three-quarters (74.3%) of the respondents reported that they do not have frequent “person-to-person” collaboration with other repository managers, and yet the collaborative environment fostered by CAIRSS is mentioned frequently in the “CAIRSS Feedback” section of being as high value to repository managers. This suggests that while repository managers may not initiate frequent collaboration between each other, the facilitative role that CAIRSS plays in providing a collaborative environment provides repository managers with valuable collaboration opportunities.

CAIRSS Feedback

Feedback for CAIRSS included areas to be targeted by CAIRSS in the next 12 months, and areas of improvement for the service. Desired target areas indicated by the respondents included the following:

- Collaborating within the community and advising on preservation strategies around both institutional repositories and their content
- Copyright workflows and policies
- Liaison with ARC regarding ERA 2012
- Assistance with ANDS-related activities, specifically the impact and importance of RIF-CS
- Building the profile of the research repository within the institution.

Areas of improvement not already mentioned in the survey included the following:

- A request for more information about what others are doing (further facilitating communication within the Community)
- Promotion of repositories
- Open Access and a possible Open Access Week campaign
- Further communication around ANDS and data management related activities

Observations

Observation 1
Results from the 2011 survey indicate an increase in the percentage of Open Access content in research repositories from 2010. With the upcoming ERA 2012 activities, it is possible that Open Access may again recede as a priority. The shift in focus in 2011 back to recruitment of Open Access content has also been commented on in general feedback to CAIRSS, and should therefore remain an area of activity for CAIRSS.

Observation 2

Responses regarding the use of OAKList indicate that use of the list is increasing slightly. It is recommended that CAIRSS continue monitoring the future of the OAKList, and provide these figures to QUT to assist in planning.

Observation 3

The apparent lack of prominent long-term preservation strategies within institutions should be something which CAIRSS takes a role in bringing to the fore of discussions within the CAIRSS Community because of a) the importance of ensuring continuity of institutional repositories and associated systems and b) the ongoing expectations of repositories to be utilized in various reporting activities for institutions.

Observation 4

The use of research repository statistics remains an area of marginal interest to the community. It is recommended that this be an ongoing area for CAIRSS to monitor and communicate developments to the CAIRSS community.

Observation 5

The area of data management continues to emerge as an area of interest to the community and should remain an area for CAIRSS to maintain a watching brief on. It will also remain important for CAIRSS to maintain a relationship with ANDS as the primary data management activity within Australia.

Observation 6

Feedback received from respondents about the amount of contact they have with other repository staff, along with general CAIRSS feedback, indicates that the facilitation role CAIRSS plays within the community remains very important. It is recommended that CAIRSS remains focused on this facilitation role within the community.

For more information / queries about the CAIRSS 2011 Repository Managers Survey, please contact CAIRSS Repository Program Coordinator, Caroline Drury at cairss@caul.edu.au