ICT in the Research Workflow Survey Report

Report authors: Janice Chan, Peter Green, Andrew Buttsworth, David Pedrotti

Date: 9 October 2019

Table of Contents

Executive Summary

Curtin University conducted a survey in 2012 to capture the ICT needs of researchers from all fields of research. The survey was a collaboration between the University Library, Research Office at Curtin, and Digital and Technology Solutions. In 2019 the eResearch Special Interest Group decided to run the survey again, to compare the current state of ICT needs of researchers with 2012 and inform University planning and support for researchers.

The overall response rate was 9% representing a good cross section of research staff and students. The outcomes of the survey are detailed in this report.

Gathering research data

- Most researchers collect their research data manually, but a notable proportion of researchers are collecting data from computer simulation and sensors, which warrants the need for automated data capture. Spreadsheets, PDF, text documents remain the main format of research data. Half of the respondents have research data in multimedia format.
- A majority of respondents (54%) requires storage space between 5GB and 500GB when working their projects. 10% of respondents need more than 12.5TB of storage space. 47% of respondents require less than 100GB of storage, a significant but expected shift from 71% in 2012.
- While most researchers store their data in more than two locations and 59% of them use the Curtin Research Data Storage (R drive) 19% of respondents store their data in only one location. Flash drives and portable hard drives remain popular storage solutions during a research project. Of the 50% of respondents who use external storage providers, most indicated that they use Dropbox. This year's results saw a decline of usage of other Curtin network drives.
- Similar to 2012, ability to collaborate with colleagues, and remote access to the data are still the most important requirements for data storage systems. Respondents clearly indicated the R Drive falls short in providing both requirements, as external collaborators cannot access the R Drive and remote connection is too slow.
- Majority of respondents collaborate with other researchers at Curtin University, over half of them also work with researchers from other Australian Universities and international collaborators. More respondents have more international collaborators than in 2012.

Processing research data

- 40% of respondents consider themselves to be undertaking computational intensive processing on their research data. 90% of respondents uses their personal computers for processing, some of them also use external research infrastructure (16%) such as Pawsey Supercomputing Centre, computer clusters at Curtin (15%) and dedicated servers (14%).
- 82% of respondents are satisfied with their current computational infrastructure. Respondents who did not find their current solutions meeting requirements attributed it to long processing time (64%), not enough space (42%), and difficult to access (25%).

Retaining Research Data

- Most respondents want to retain data for themselves and others. Nearly two-third of them want their data citable, but only one-third want to make their data publicly discoverable. Staff are slightly more likely to make their data publicly discoverable than students.

- The Majority of respondents want to retain their data between 5 and 10 years. 22% of respondents indicated that they want to retain their data for less than 5 years. It is worth noting that the Western Australian University Sector Disposal Authority's minimum period for retention of research data is 7 years after the date of publication or project completion. Responses from HDR students shows a tendency to not retain their data for the minimum retention period.
- Most of the respondents (82%) indicated that their storage needs after project completion are the same or similar to their storage needs during the research project. This year's results mirrored that of 2012. Providing authorised access to data and high security for confidential and sensitive data are still the most important features of a data retention storage system.

Services feedback

- Staff are more likely to use the R Drive than HDR students. 45% of respondents who have used the R Drive commented positively about their experiences. Common themes of issues encountered by researchers are the lack of access control by project lead, application to access is complicated, slowness when connecting remotely, and general concerns about backup.
- Only 5% of respondents have used the Curtin data publication service and commented positively about it. 60% of those who have not used the service are not aware of it.

ICT needs of researchers who are HDR students and those who are staff are very similar. It is suggested that it would be desirable to provide students the same access to infrastructure and support as staff. The R Drive is well utilised and fit for purpose, but the limitations on access by external collaborators and remote access speed needs to be addressed. The survey results indicated that support and guidance on data management best practice, and connecting researchers to services available to them is crucial to maximise the use of resources and enhance researchers' experience.

Introduction

Curtin University conducted a survey in 2012 to capture the ICT needs of Curtin University researchers from all fields of research. The Survey report: ICT in the Research Workflow from 2012 can be accessed via espace http://hdl.handle.net/20.500.11937/17497.

In 2019 the eResearch Special Interest Group ran the survey again. In order to compare researchers' needs and how they have evolved, the survey questions remain largely unchanged.

To quote the 2012 report, the survey was formulated around three key concepts:

1. Focus on capability, not technology.
2. Focus on gathering existing requirements, not future desires.
3. Utilise a generic research workflow that researchers could easily relate to.

New questions seeking feedback on the Research Data Storage (R Drive) and Data Publication Service were added to the survey. These are services that did not exist in 2012 but are now a substantial part of the research ecosystem at Curtin University.

The survey covers the following key topics:

1. Gathering research data (3 questions)
2. Working with research data (5 questions)
3. Processing research data (4 questions)
4. Retaining research data (4 questions)
5. Research data services feedback (7 questions)
6. Demographic (4 questions)
7. Further feedback (3 questions)

The survey was conducted from 17 June 2019 to 8 July 2019 and comprises 30 questions, all of them optional. An email invitation to participate was sent to 1,099 active continuing staff who have 'Research Only' or 'Teaching and Research' roles; and 1,930 higher degree by research (HDR) students with course status 'Admitted'. 262 researchers started the survey, 191 (73%) completed it. The response rate of HDR students was half that of staff. The overall response rate was 9%. The average time taken to complete the survey was approximately 15 minutes.

Out of the 191 responses, 99 (52%) respondents were HDR students. The majority of the respondents were from the Faculty of Science and Engineering (39%), followed by Health Sciences (27%), Humanities (18%), and Business and Law (16%). This reflects a good cross section of the research staff and students of Curtin University.

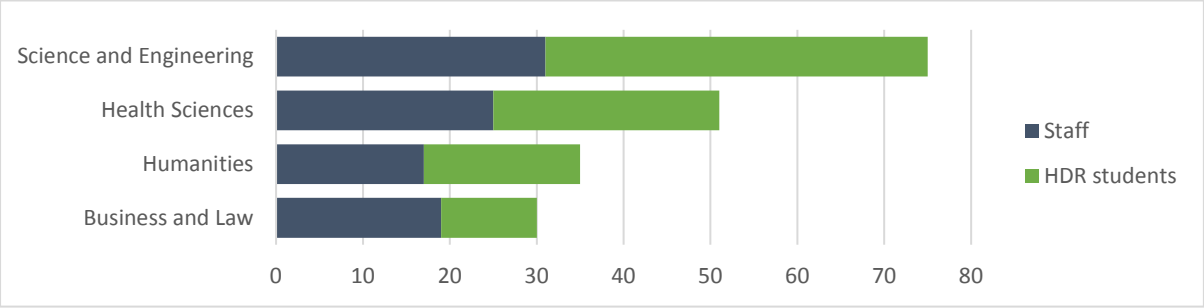

Figure 1 Respondents by faculty and role

Question 27 asked respondents to list their Field of Research (FoR) codes. The chart below shows the distribution of research areas by FoR codes at 2-digit level.

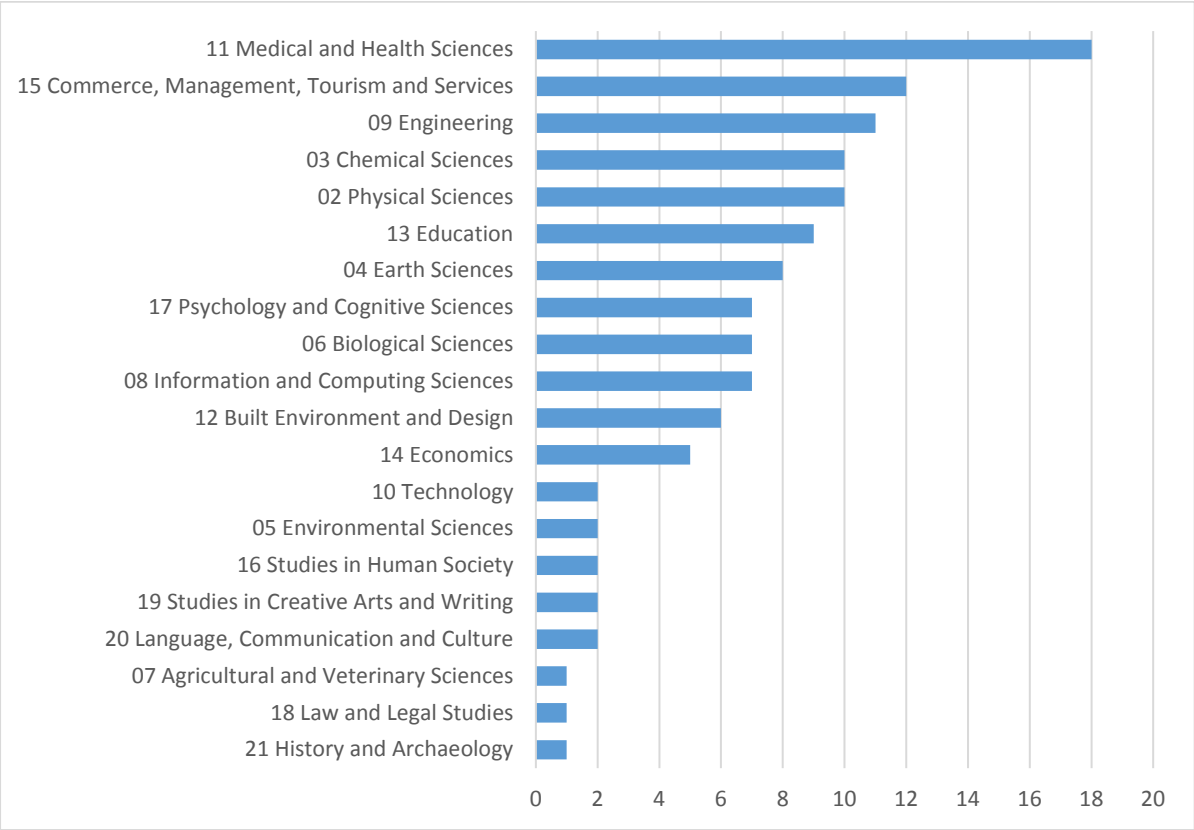

Figure 2 Count of valid FOR codes provided in Q27

This report covers results and analytical narratives from both completed and partially completed surveys. Where there is notable difference in the responses between HDR students and staff, breakdown of responses by role are included. This report also includes findings from the 2012 report for comparison where applicable.

Gathering research data

Source of research data

Question: What is the source of your research data?

222 respondents answered this question, respondents may select more than one option. 407 selections were recorded.

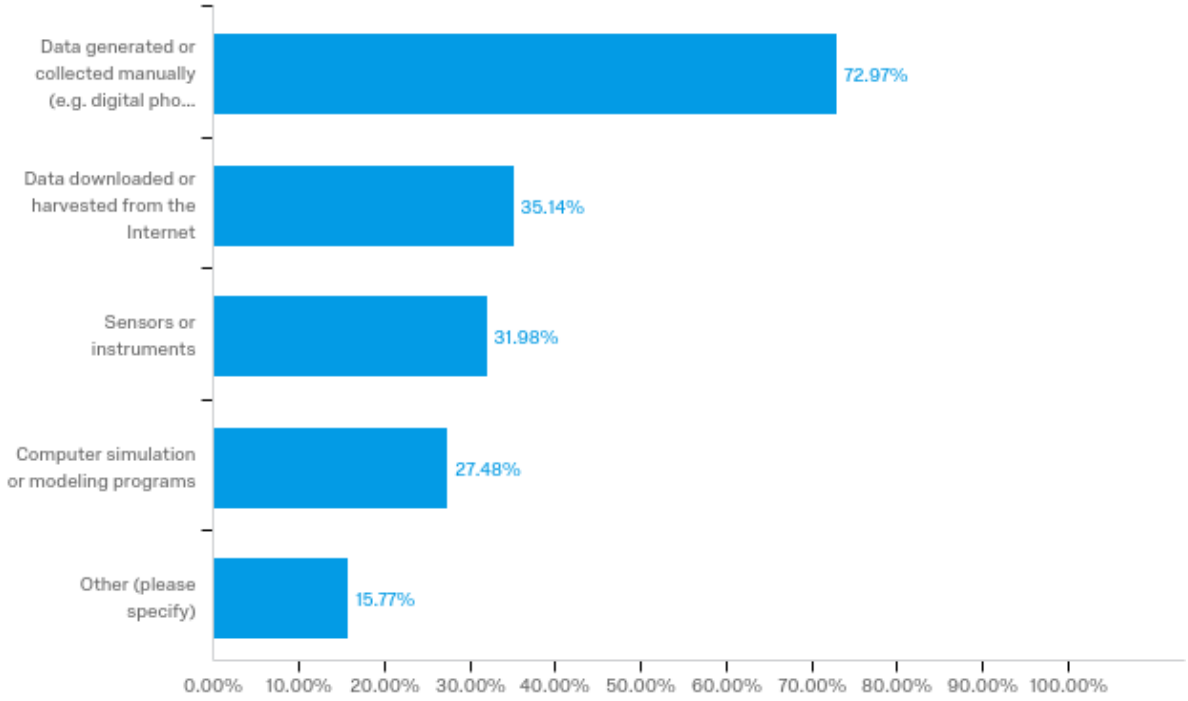

Figure 3 Source of research data (% of responses)

73% of respondents collected their research data manually, follow by a fairly equal distribution between downloaded or harvested from the Internet (35%), sensors or instruments (32%) and computer simulation or modelling programs (27%). Computer simulated and sensors data usually produce sufficient data to warrant the need for automated data capture and ingest. Automated data capture and ingest into research data storage could simplify research workflow and enhance user experience.

25% of respondents gather research data from one source, 75% from more than one source, 39% from 3 sources or more. This is a notable shift from 2012, where only 17% of respondents gathered data from 3 sources or more.

33 out of 35 respondents who selected “other” provided further comments. Most of the comments could be classified as “Data generated or collected manually”, such as interviews (14), focus groups (4), surveys (2), data generated from experiments (2). Other comments included:

- Fieldwork
- Health data and images
- Health data collected for administrative purposes
- Data from government systems
- Data from Landgate

- Archives, transcripts
- Purpose built apps and online applications

Data format

Question: Which of the following categories apply to your research data?

222 respondents answered this question, respondents may select more than one option. 408 selections were recorded.

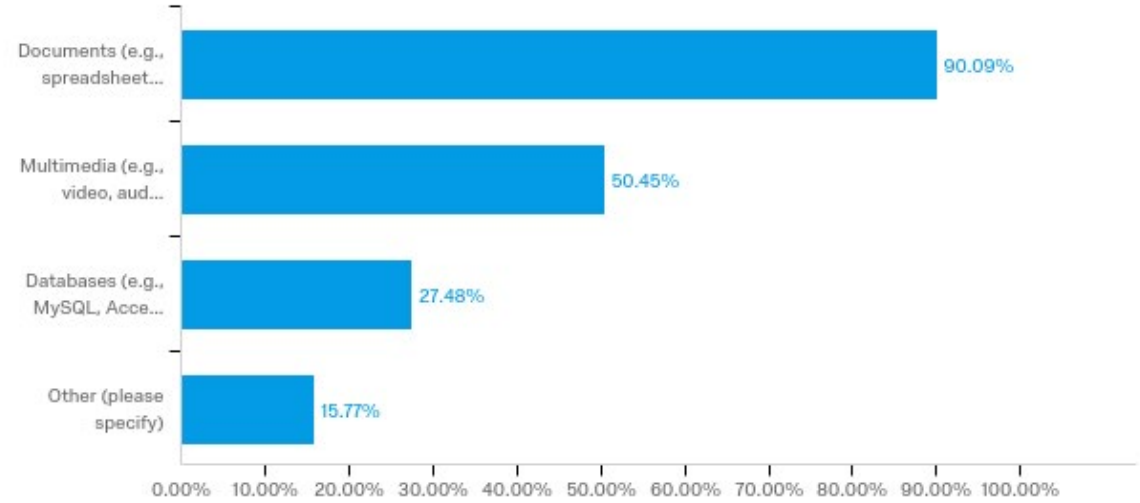

Figure 4 Types of research data (% of responses)

90% of respondents have data in the form of documents, 50% in multimedia format such as video, audio, and images, and 27% in databases. These results are very similar to the 2012 findings.

36% of respondents have only one form of data, of which 75% are in document format.

Spreadsheets, PDF, text documents remain the main format of research data.

Text responses provided by respondents that selected “Others” that are not classified in the three choices above tended to be domain-specific or proprietary data structure and output files included:

- C3D
- DNA sequence
- Binary files
- Radio astronomy data
- Molecular structure data
- Maps
- CPLEX
- FCS
- Experiments
- GIS systems (QGIS, ArcGIS)

Databases and software

Question: Which of the following database software solutions do you use?

135 respondents answered this question, respondents may select more than one option. 164 selections were recorded.

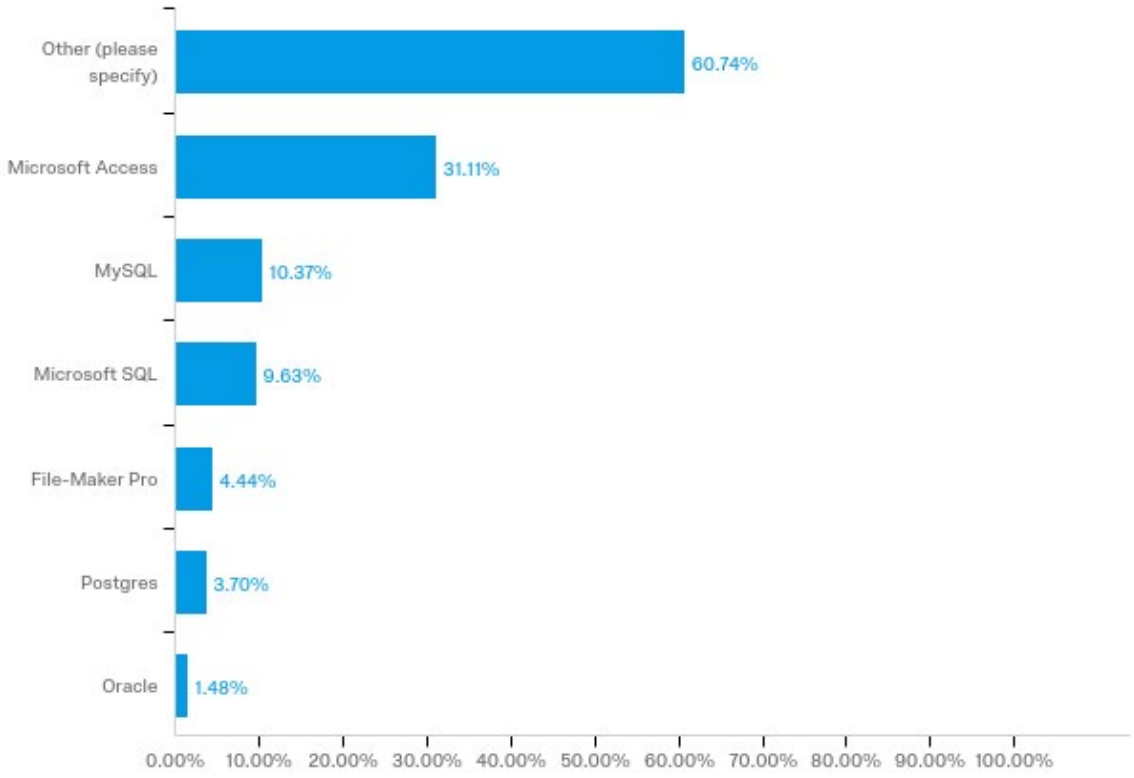

Figure 5 Databases and software (% of responses)

Down from 40% in 2012 to 31% in 2019, Microsoft Access is still the most popular database solution. The order of preference is the same as the results from 2012, with the percentage for responses for each solution much lower in 2019 results.

However, 15% more respondents (61%) have answered “Other” in this survey. The most common solutions provided are, Nvivo (16); SPSS (15); Excel (14); Stata (5); SQLite (4); MATLAB (3). Other database and software solutions include:

- Breeders assistant
- Topcat, ADQL
- MongoDB
- Stata
- MPlus
- Redcap
- Katmandoo
- Qualtrics
- CFD Ansys
- By a database IT person
- Final Cut Pro
- Adobe Premiere
- QPS QINSy
- GenStat

- DataStream
- AlphaAnywhere
- PRISM Graphpad
- R

Working with Research data

Data storage space

Question: How much storage space do you need to work on your research project?

204 respondents answered this question.

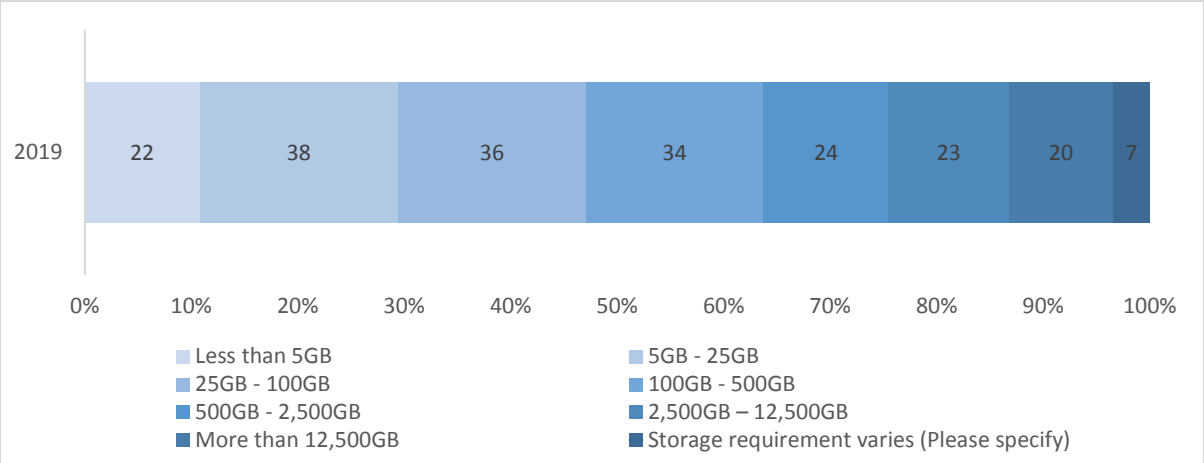

Figure 6 Storage space for data during project (2019)

Majority of respondents require 5-25GB (19%), followed by 25-100GB (18%), 100-500GB (17%). 47% of respondents need less than 100GB of storage space during their project. 10% of respondents need more than 12.5TB of storage space.

4 out of the 7 respondents who answered “storage requirement varies” provided further comments. Most of them answered “not sure” or “depends”, one respondent commented “Datasets range from <1GB to >20TB”.

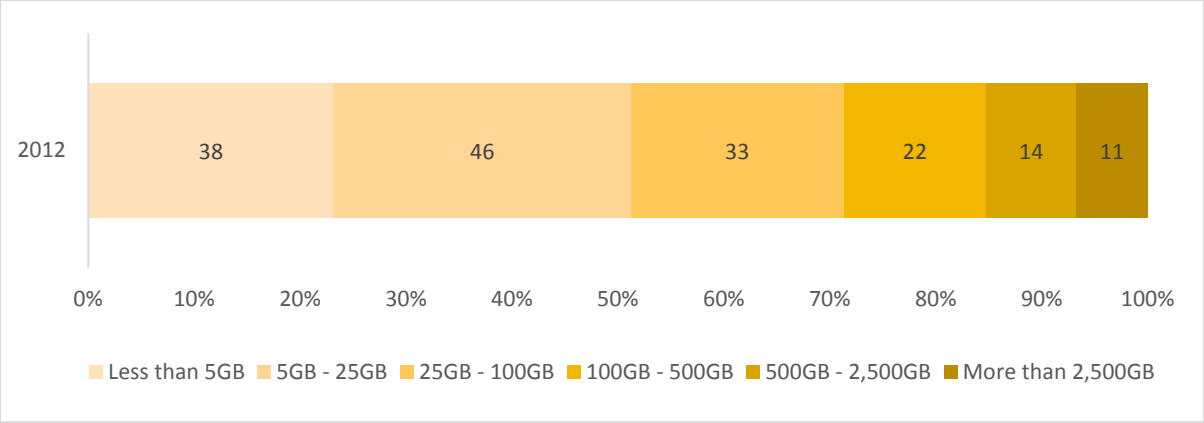

Figure 7 Storage space for data during project (2012)

Comparing the results from the two surveys, we saw storage space needs during a project has significantly increased in the last 7 years. In 2012, 71% of respondents required less than 100GB storage. Only 7% of respondent required more than 2.5GB of storage space data during project, while in 2019, this has gone up to 21%. This aligns with growth and availability of data and storage.

Data storage locations

Question: While working, where do you store your research data?

204 respondents answered this question, respondents may select more than one option. 522 selections were recorded.

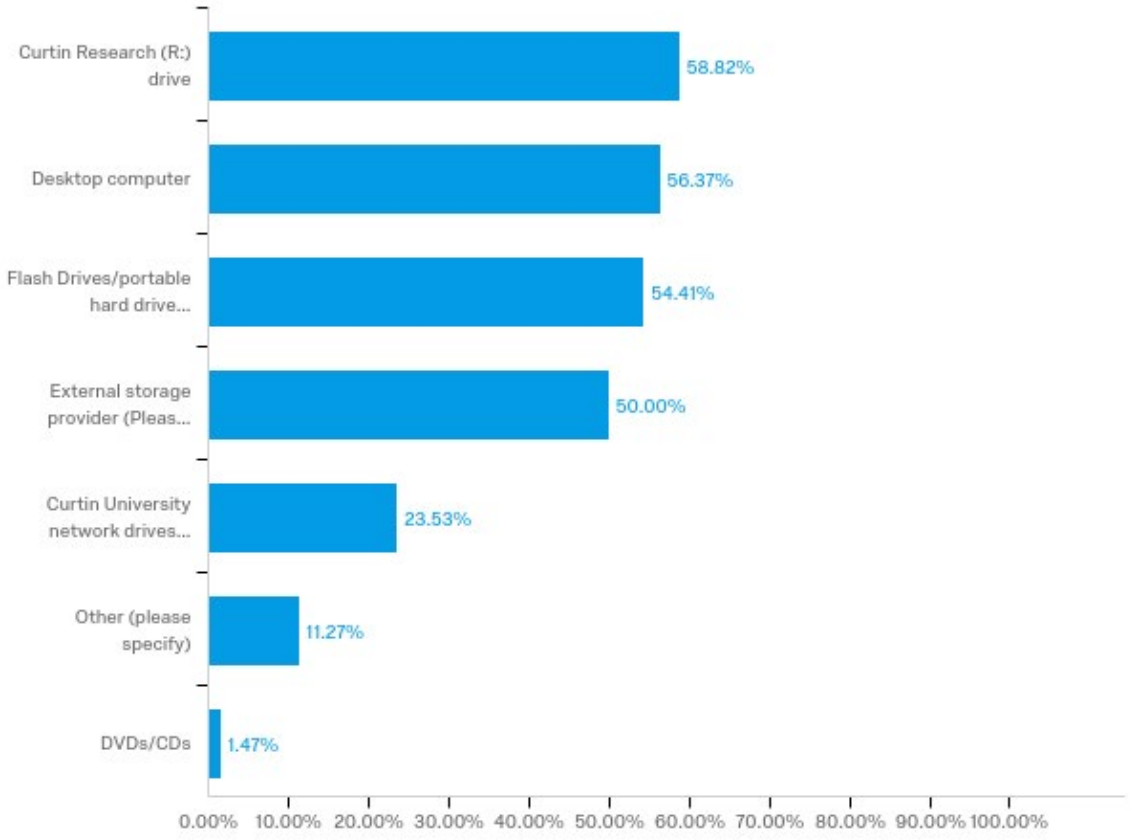

Figure 8 Current data storage locations (% of responses)

81% of respondents store their data in more than two locations, 50% in more than three locations.

59% of respondents store their data in Curtin's R Drive, and 24% on other Curtin network drives. Desktop computer (56%), flash drives and portable hard drives (54%) remain popular storage solutions for researchers while working on their research data.

50% of respondents store their data with an external storage provider. Of those who provided further comment, Dropbox (31) was the most mentioned provider, followed by Google Drive (15), OneDrive (14), Cloudstor (8), and Pawsey (4).

For respondents who selected 'Other', comments included:

- Pawsey
- Local NAS
- Various internal department or lab shared network drives

- State government server
- Hardcopy

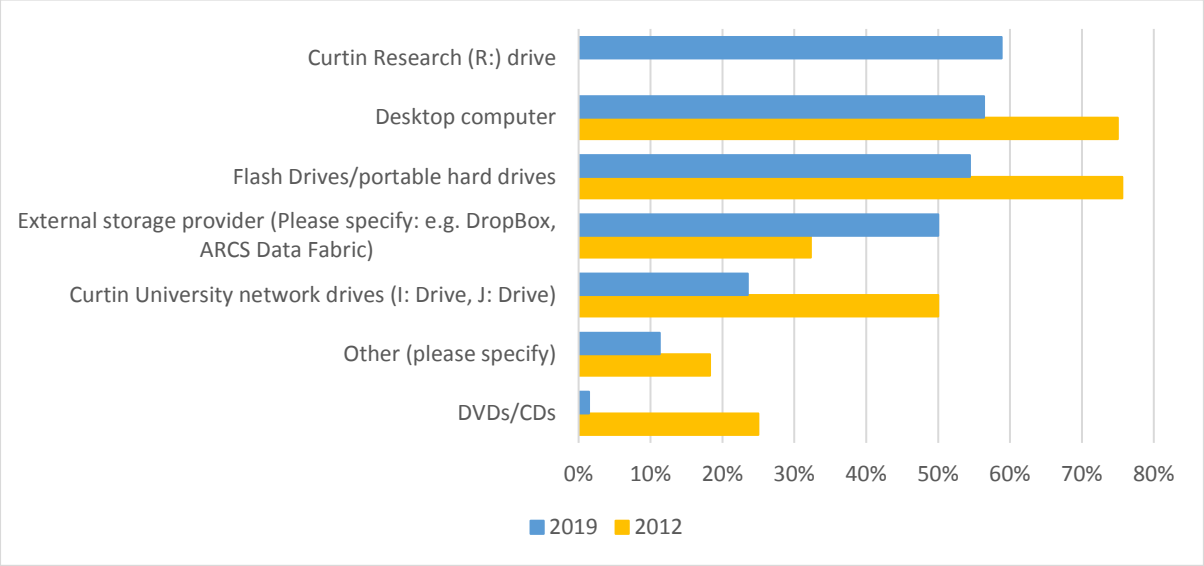

Figure 9 Data storage locations (% of responses) comparison

Figure 9 shows the comparison of data storage options between 2012 and 2019. With the introduction of the R Drive, there is a significant decrease of use of other Curtin network drives, and moderate decline for desktop computers, flash drives and portable hard drives, although they are still the most common storage solutions. 2019 results also saw an increased usage of external storage provider. DVDs/CDs are almost phased out.

Requirements for working with research data

Question: What special requirements do you have for working with this research data?

203 respondents answered this question, respondents may select more than one option. 485 selections were recorded.

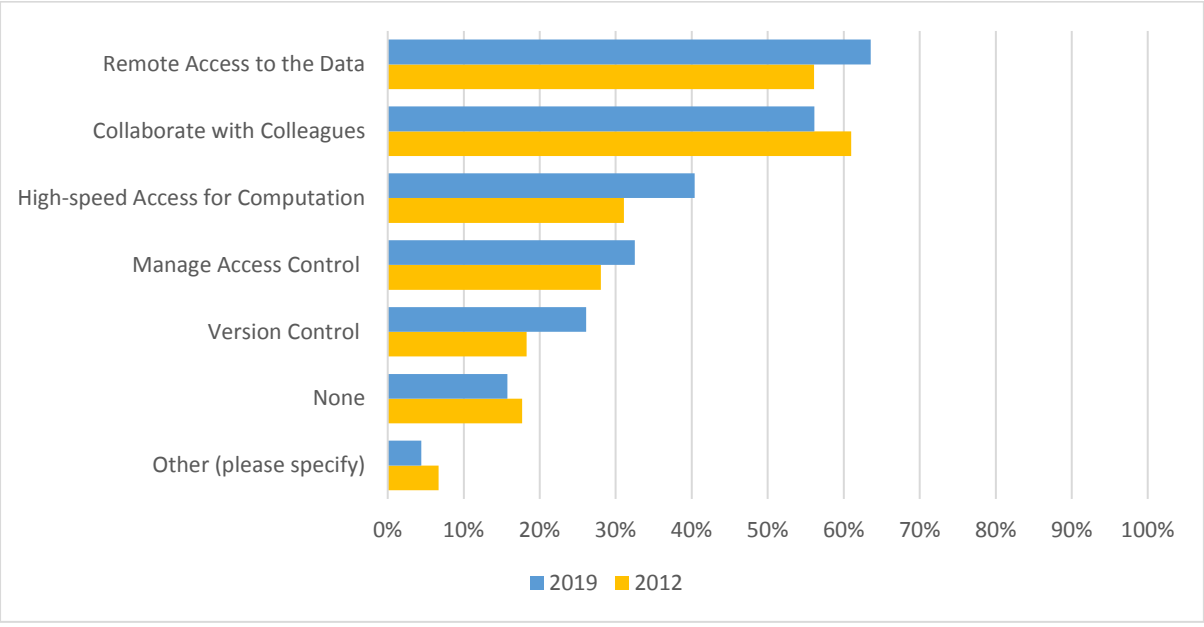

Figure 10 Requirements for working with research data (% of responses)

Researchers' requirements for working with research data in 2019 are very similar to 2012. Ability to collaborate with colleagues and remote access to the data are still the most important requirements for data storage systems.

There is a notable increase for high-speed access for computation, version control, and remote access to the data.

Comments for 'other' included:

- Collaboration with people outside Curtin
- Blockchain
- Encryption
- Powerful graphical processing unit
- Confidentiality
- VR/Panoramic editing and screening systems

Access to research data

Question: Where do you need to be able to access your working research data from?

202 respondents answered this question, respondents may select more than one option. 557 selections were recorded.

Majority of respondents require access to data from home (91%), any Curtin campus (76%), and while traveling (69%). This is similar to the 2012 result. A decline in the need to access from remote research sites is noted but the response rate for this question in 2012 was low and the decline may not be significant.

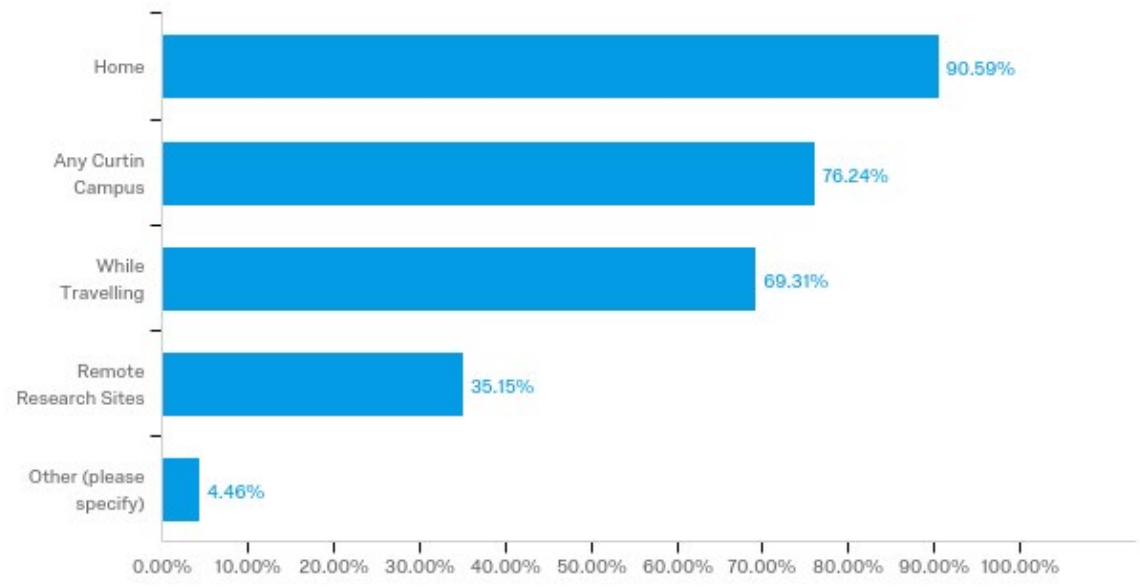

Figure 11 Remote access to data (% of responses)

Comments for 'other' included:

- Other universities
- Rural areas
- Health sites

- Collaborator sites/offices
- Low bandwidth third world site

Collaboration

Question: Who do you collaborate with?

196 respondents answered this question, respondents may select more than one option. 515 selections were recorded.

Majority of respondents collaborate with other researchers at Curtin University, followed by international collaborators, researchers at Australian Universities, industry partners, Researchers at State or Federal Government Agencies (e.g. CSIRO, Water Corporation etc.), and researchers at medical research centres.

7 respondents answered 'Other', but all of the answers can be categorised under one of the choices available.

More respondents have international collaborators than in 2012. International collaborators was ranked 4th by percentage of responses from the 2012 results, behind researchers at Curtin, researchers at Australian Universities, and industry partners.

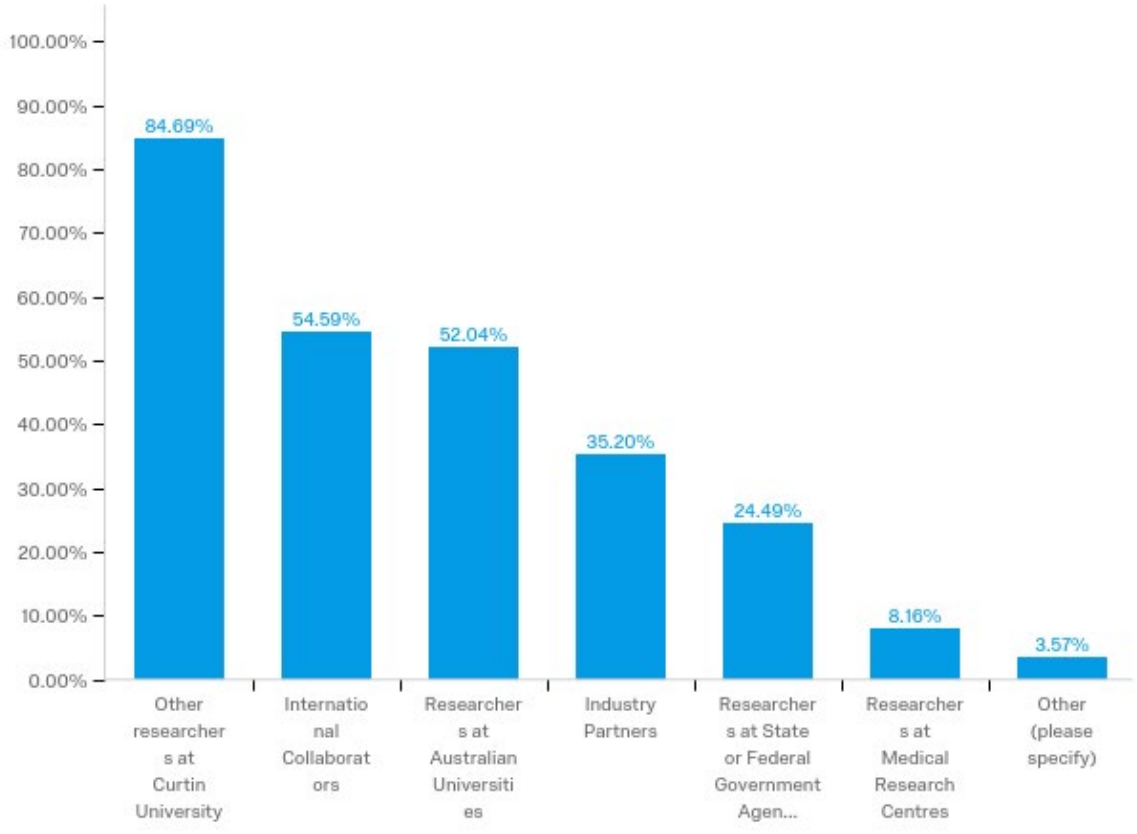

Figure 12 Collaboration (% of responses)

The below graph shows the breakdown of collaborators by how respondents answered question 26, "Are you an HDR student?". HDR students are much less likely to have external collaborators, in particular researchers from government agencies and medical research centres.

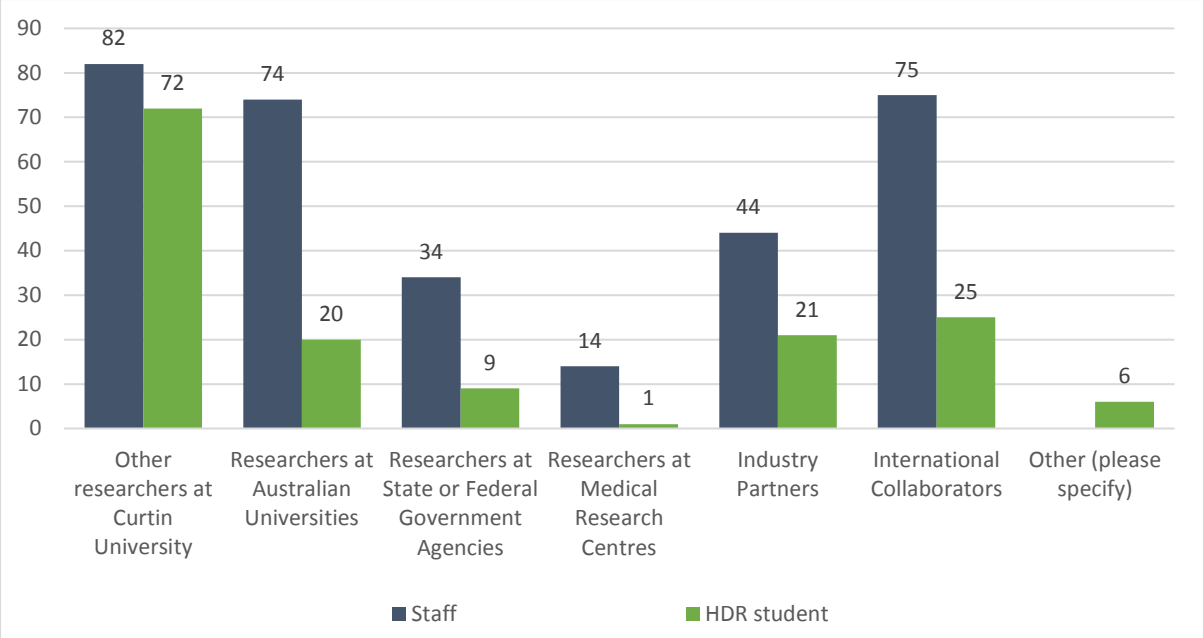

Figure 13 Collaboration by researcher type

Processing Research Data

Computational infrastructure

Question: Do you use computationally intensive processes on your research data?

201 respondents answered this question. 40% (81) of respondents perform what they consider to be computational intensive processing on their research data. An increase compared to 34% (56) from the 2012 results.

Question: What infrastructure do you use to process your research data?

196 respondents answered this question, respondents may select more than one option. 274 selections were recorded.

123 (63%) of respondents use only their personal computers to perform computationally intensive processing. It is unclear how respondents define computationally intensive processing.

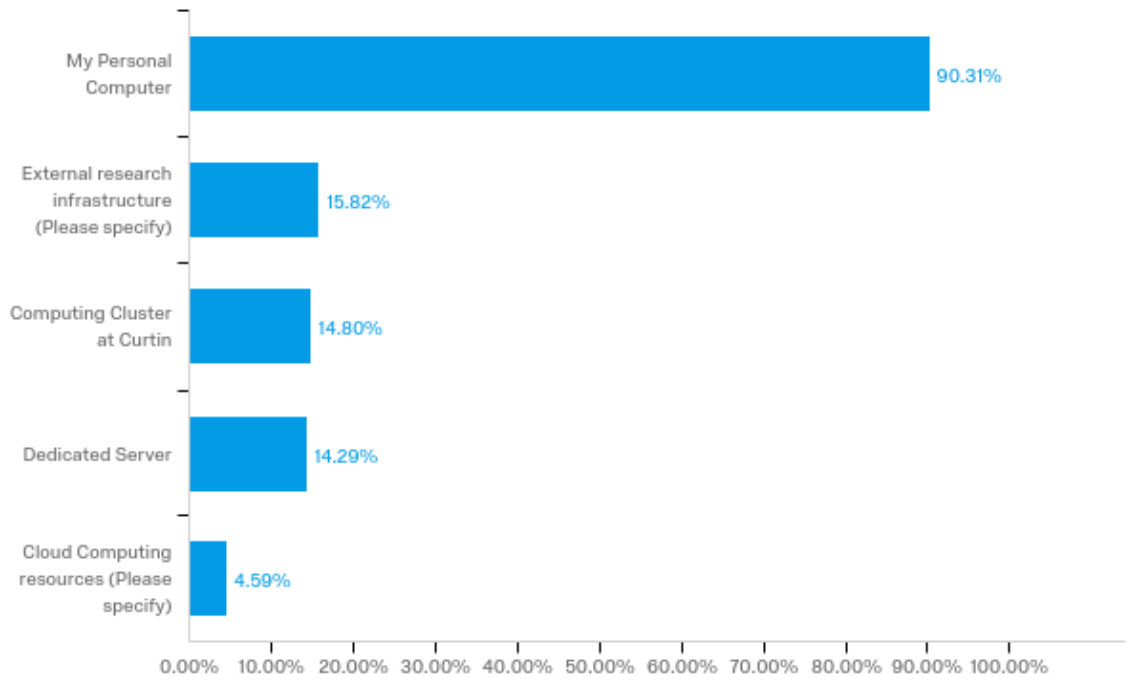

Figure 14 Processing infrastructure (% of responses)

The use of external research infrastructure, computing cluster at Curtin, and dedicated server are quite evenly distributed.

31 respondents selected “External research infrastructure”, 18 of them specified that they use Pawsey Supercomputing Centre. The rest included Secure Unified Research Environment (SURE), AWS, Nectar, Australian Bureau of Statistics data lab, WA Department of Health, and via database IT personnel.

5 respondents selected “Cloud Computing resources”, 3 specified that they use Pawsey, 1 AWS, and 1 via a database IT personnel.

Adequacy of current computational infrastructure

Question: Does your current solution meet your processing requirements?

200 respondents answered this question.

82% (164) are satisfied with their current computational infrastructure.

If respondents selected “No”, they were asked to specify why it does not meet their processing requirements. 36 respondents answered this question, respondents may select more than one option. 57 selections were recorded.

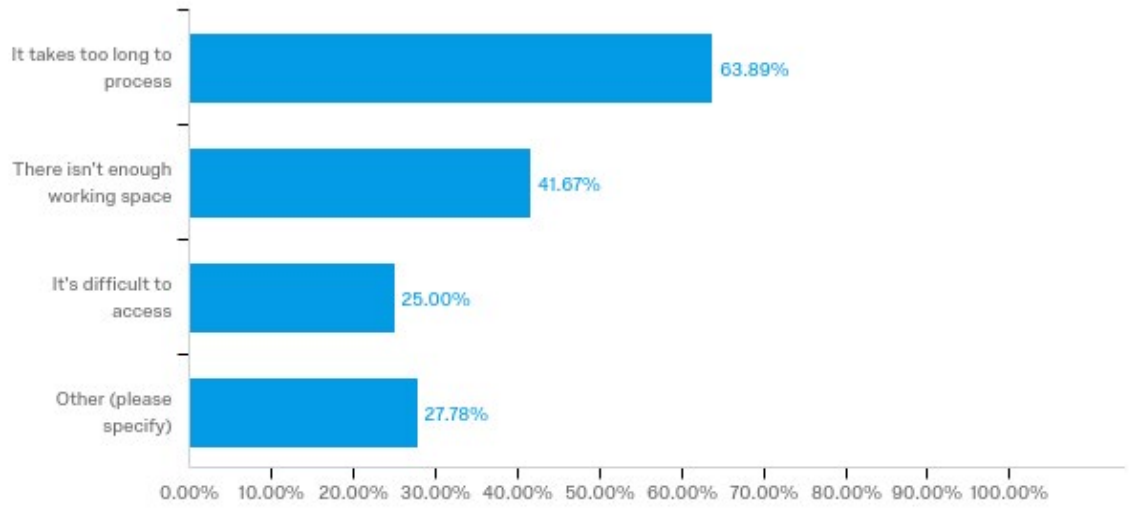

Figure 15 Current processing challenges (% of responses)

Overall, the main challenge is it takes too long to process, followed by not enough working space. Breaking down the responses by infrastructure used from the previous question, it seems these challenges were mostly related to personal computers.

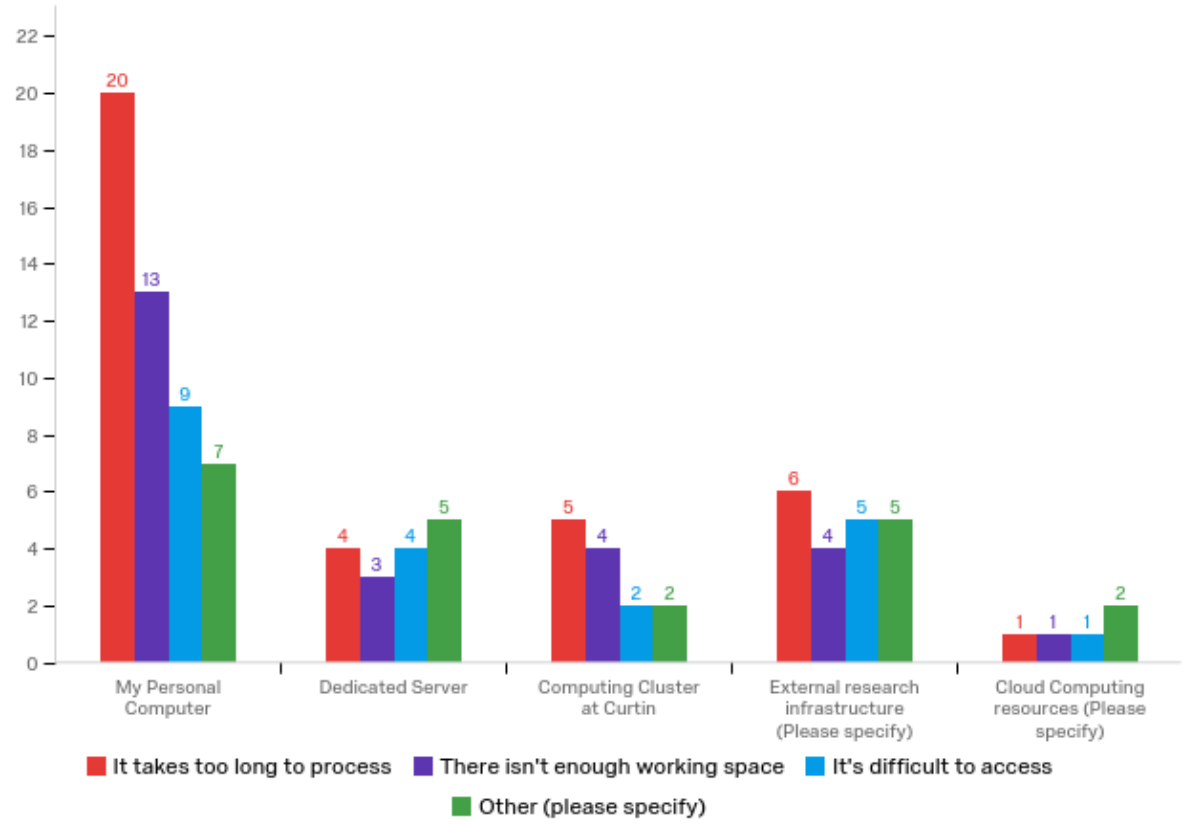

Figure 16 Current processing challenges breakout by infrastructure used

10 respondents selected “Other”. It is worth noting that 6 out of 7 respondents who provided additional comments indicated that they use dedicated server, computing clusters, external research infrastructure and cloud computing resources. Comments included:

- *It's not reliable - lots of down time*

- *Too much time lost due to technical issues and downtime*
- *Pawsey utilisation is high*
- *difficult to share*
- *Shared resources and waiting queue*
- *We have changing requirements*
- *I need more technical and statistical support than is available remotely than I get*

While the percentage of responses to processing challenges are similar to the 2012 result, the comments from 2012 were more about lack of collaborative workspace, knowledge, high level of security, and not enough computing resources in terms of RAM and software – likely referring to personal computers. This year, only one respondents who provided the above comments uses mainly their personal computer, so the comments are mostly related to usage of dedicated servers, clusters and HPCs.

Retaining Research Data

Reasons for data retention

Question: At the end of your project, what do you need to do with the data that supports your research result?

195 respondents answered this question, respondents may select more than one option. 386 selections were recorded.

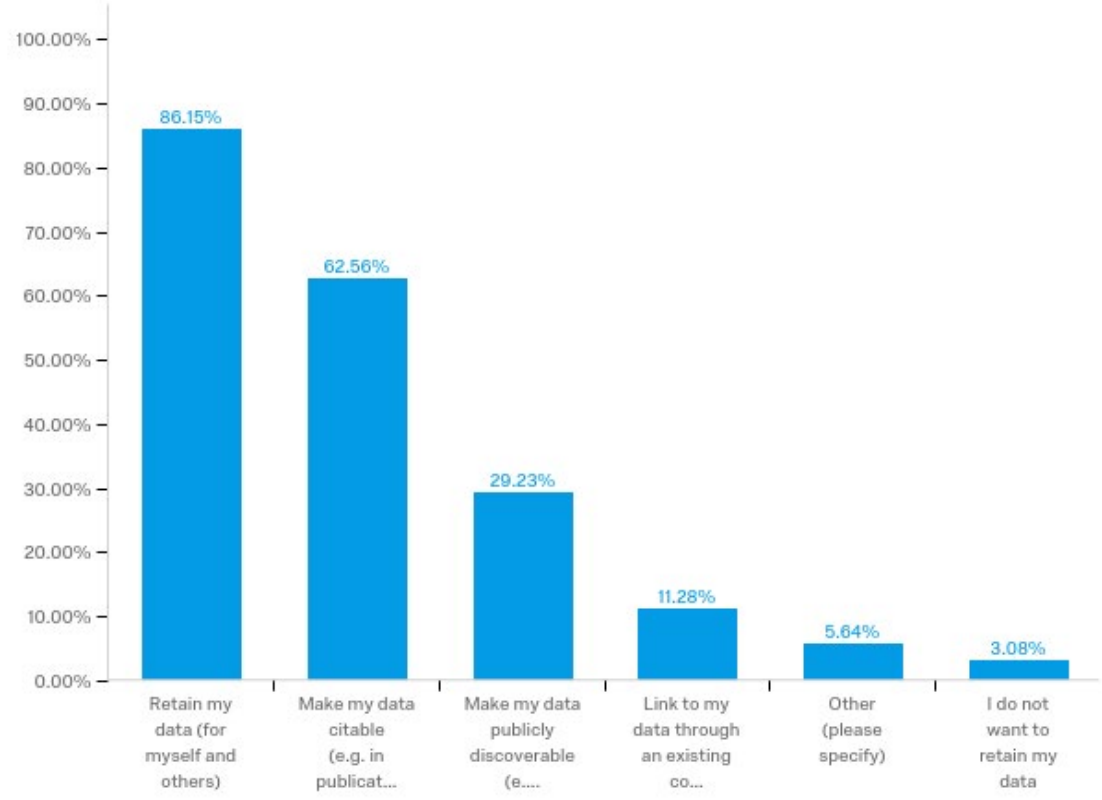

Figure 17 What to do with data after research project (% of responses)

86% of respondents want to retain data for themselves and others. 63% want to make their data citable, but only 29% want to make their data publicly discoverable. These results are very similar to the 2012 survey, except in this survey, 3% (6) of respondents do not want to retain their data, but in 2012, 0.6% (1) respondent selected that answer.

Comments for "Other" included:

- Deposit to open data repositories
- Handover data to the state government
- Data needs to be destroyed within a certain timeframe
- Maintain strict data control as per ethics clearance
- Edited into various film projects for semi-public screening
- Not sure of the implications of the various options yet
- Data is confidential
- Archive

Staff are slightly more likely to retain their data and make their data publicly discoverable. About the same number of HDR students (3) and staff (2) responded they do not want to retain their data.

Data retention period

Question: How long do you need the data that supports your research result to be retained for?

194 respondents answered this question.

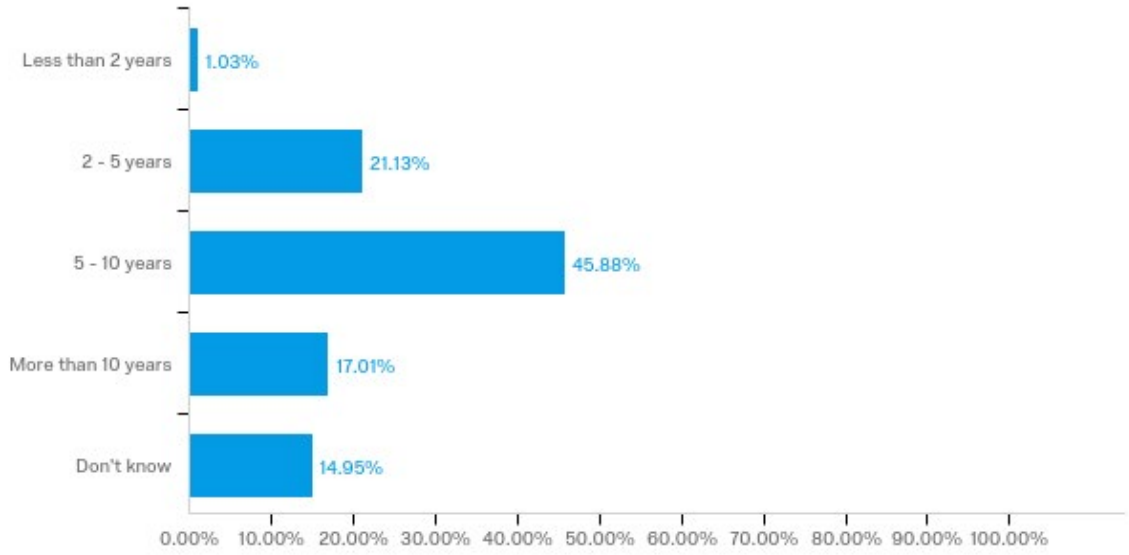

Figure 18 Retention period (% choice count)

The question captures the respondents' personal appreciation of data retention periods. It is worth noting that according to Curtin's Information Management and Archives Storing Research Data advice sheet, the minimum period for retention of research data is 7 years from the date of publication or project completion

20% of respondents would retain the data for 2-5 years, majority of the respondent would retain their data for 5-10 years. 15% (29) of respondents do not know how long they need to retain their research data. These are very similar to the 2012 results.

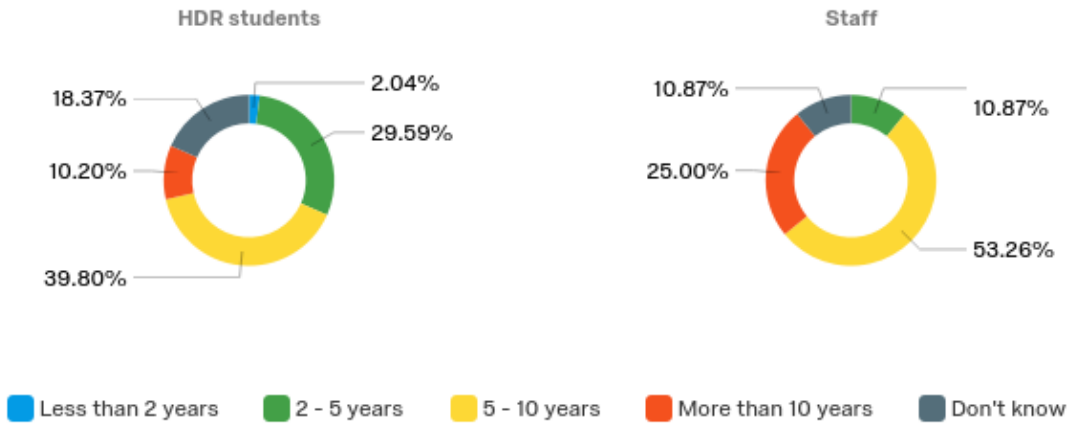

Figure 19 Retention period (breakdown by role)

Only HDR students want to retain their data for less than 2 years. HDR students are much more likely to retain their data for 2-5 years. 18% of respondents who are HDR students and 11% who are staff do not know how long they need to retain their data.

Long term storing space

Question: How much storage space do you need for the data that supports your research result?

194 respondents answered this question.

21% of respondents stated they require the same amount as their original working research data. For respondents who have different storage requirements after project completion, they may indicate their storage needs by the same size range as per the Data Storage Space question.

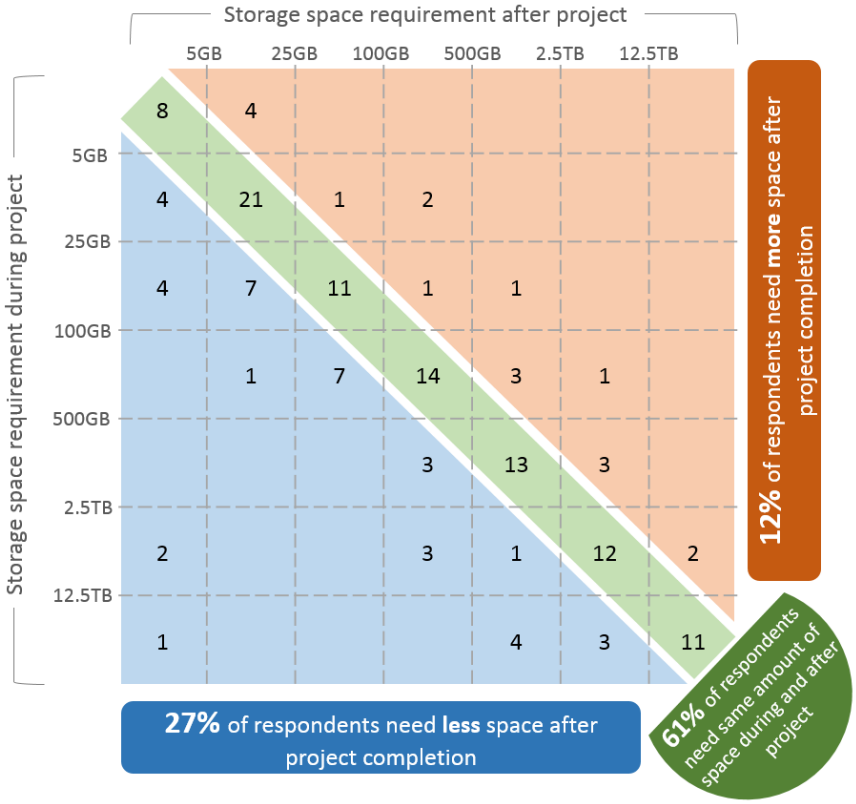

Figure 20 Change in storage space during and after project

Figure 20 compares respondents' storage needs during and after projects by mapping individual respondent's selection for both questions. The numbers on the grid represent the total number of respondents who selected that particular size range combination during and after project. E.g. 1 (bottom left corner) respondent indicated that they require more than 12.5TB of storage during project and less than 5GB after project completion.

Although most respondents require different amount of storage after a project is completed, the majority of their storage requirement did not change significantly. As shown above, 61% of respondents have storage space requirement that falls within the same size range during and after project. 27% required significantly less storage after project, and 12% required significantly more storage after project.

Features of data retention storage

Question: What are the required features of the storage system that will hold this data?

183 respondents answered this question, respondents may select more than one option. 418 selections were recorded.

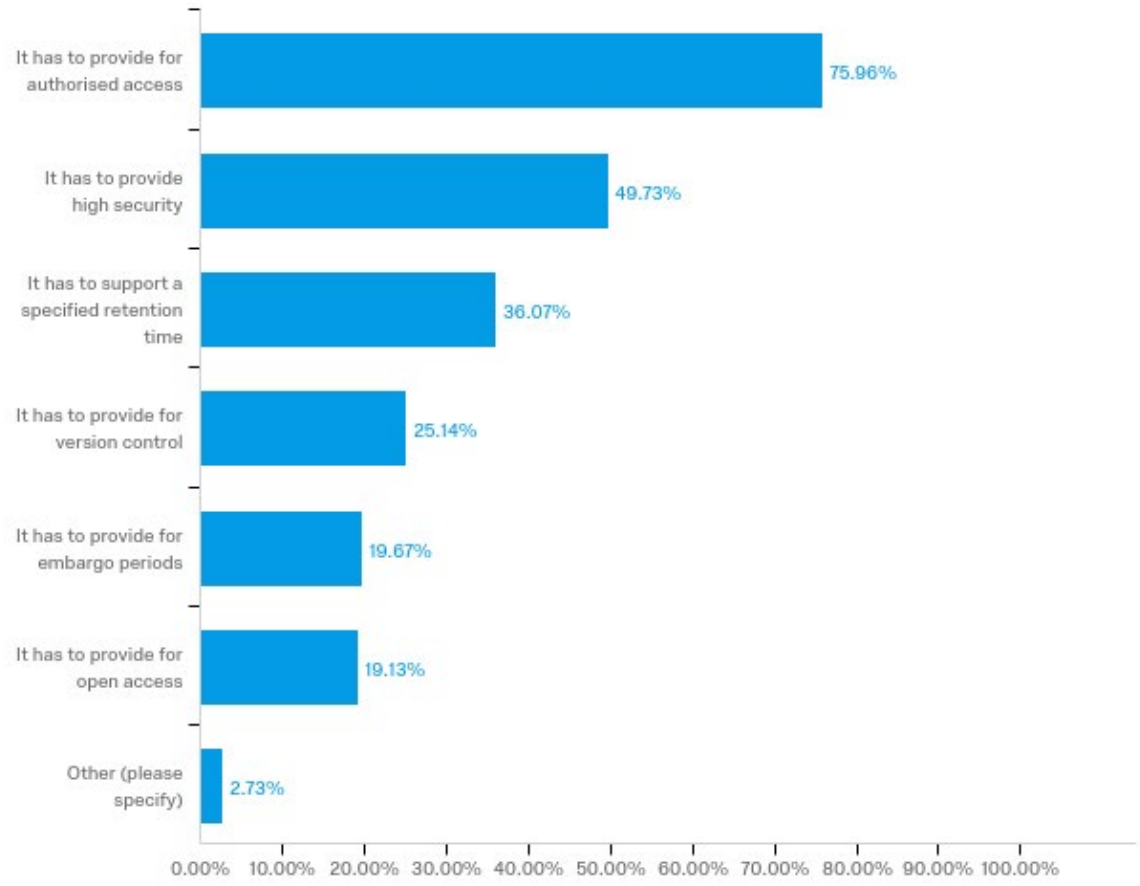

Figure 21 Features of data retention storage (% of responses)

This year's results mirrored that of 2012. Providing authorised access to data and high security for confidential and sensitive data are still the most important features of a data retention storage system.

The 2012 report foresaw the requirement for providing open access to research data to rise, however, the requirement has not increased despite an increasing number of publishers and funders

requiring data to be made openly available. In fact, the requirement to provide for embargo periods overtook open access, though by less than 1 percent. It is worth noting that the research community has been promoting FAIR data, rather than open data in the past few years.

Comments for "Other" included:

- Redundancy in case of hardware failure
- Important that IPA script not be corrupted
- Off-site access
- Part of research data be made publicly available

Services Feedback

This section of the survey captured respondents' feedback on service and infrastructure that were put in place after the 2012 survey.

R Drive

Question: Have you used the R Drive for storage?

193 respondents answered this question. 59% of respondents have used the R Drive. More staff have used it for research data storage.

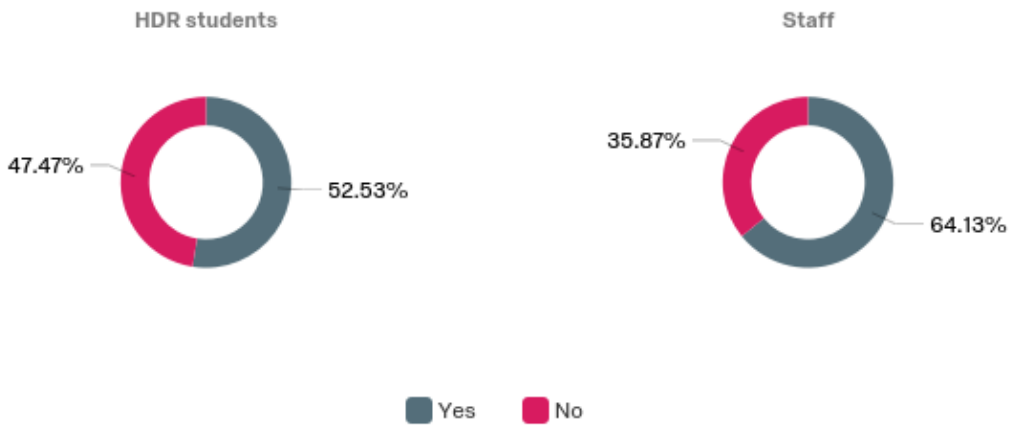

Figure 22 Have you used the R Drive for storage? (breakdown by role)

Question: Would you like to give feedback about your experience?

If respondent answered Yes to the previous question, they will be shown this question.

62 respondents commented.

45% of respondents commented positively about their experience. Most frequent words in the comments are: access (25); drive (20); data (17); slow (12); good (11).

Comments are aggregated below:

+ Supervisor has access to data
+ Works well
+ Meets requirements
+ Additional storage requests and access queries have been dealt with promptly
+ Ability to access remotely through VPN
– Application to access R Drive is complicated
– User access control process needs improving
– User access controls are not transparent
– Slow
– Unstable
– No version control
– More controls could be given to project lead
– Need to streamline process to give access to additional users to R Drive project folder
– DTS service desk staff do not seem to understand, hence provide correct advice on additional user access to existing project folder
– External collaborators cannot access the R Drive
– Access from external facilities can be slow
– Need compute facilities or command line interface
– Used as backup, not processing because it is slow if accessing remotely
– Doesn't seem to work with EndNote
– General concerns about backup of R Drive

Question: Why have you not used it?

If respondent answered No to the previous question, they will be shown this question.

66 respondents commented.

Comments are aggregated below:

- School has its own research data storage, e.g. N drive
- Seems clunky, time consuming to setup
- Linux user
- R Drive is too small
- Not available from respondent's campus
- Does not meet respondents' need
- Cannot access remotely
- Not aware it exists
- No information on how to access or use it
- Prefer other solutions
- Restrictions in terms of what can be held
- Not required
- Personal laptop storage is enough
- Was advised by DTS only need to store the data once the research is complete

The responses indicated that many researchers are not aware of the support available and not familiar with the features of R Drive, e.g., ability to access remotely via VPN. Comments from various respondents also indicated that staff might have provided incorrect advice to researchers. Self-help resources are available but it seems researchers are not aware of them. There is opportunity to improve general awareness of data management infrastructure and support available to researchers.

Curtin data publication service

Question: Have you used the Curtin data publication service?

190 respondents answered this question. Only 5% of respondents have used the Curtin data publication service, which includes describing and publishing research datasets to Research Data Australia and minting of Digital Object Identifiers, a free service provided by the Library.

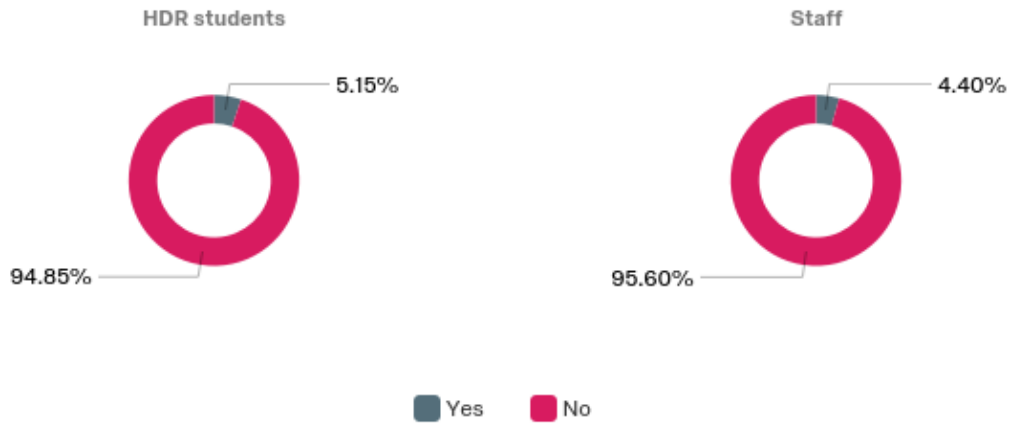

Figure 23 Have you used the Curtin data publication service? (breakdown by role)

Question: Would you like to give feedback about your experience?

If respondent answered Yes to the previous question, they will be shown this question.

3 respondents commented. 2 of them commented it worked well, one of them commented that most of their data has restrictions on who may access it.

Question: Why have you not used it?

If respondent answered No to the previous question, they will be shown this question.

131 respondents commented. It is worth noting that 60% of respondents commented that they are not aware of the data publication service.

Comments are aggregated below:

- Not aware of the service
- No need to make data publicly accessible
- Data already publicly accessible elsewhere
- Data is identifiable, confidential
- Difficult and complex when compare to alternatives
- Unclear what happen to the published data after the researcher leave Curtin
- No time
- Will use it when there is a need
- Too much to learn, just another thing
- Why would/should I use it?
- Curtin has too many systems and are generally aren't useful
- Publication in the journals that I do often requires data to be connected to the publication
- Not applicable to my qualitative data
- Do not have permission to publish photographs and archival documents

The responses indicated that many researchers are not aware of the Curtin data publication service. Some comments also showed a lack of understanding of what data can be published. More promotion, training, and awareness raising activities should be considered. A more streamlined, user-friendly data management and publication workflow might attract researchers to use the infrastructure at Curtin not just to publish open data, but ensure that Curtin's research data is FAIR.

Research infrastructure

Question: Does your research project have ICT needs that are not currently being met?

102 respondents answered this question. 62 answered "No", "not applicable", or "don't know".

Other comments included:

- Small, one-off tasks are completed as a good will gesture by CIC staff who are curious and helpful.
- Not enough parallel processing
- Lack of network connections to equipment

- Inability to install software, the change to admin right rules is annoying and slow down research
- Lack of Linux support, lack of Linux-friendly file sharing system
- Access issues present on Mac devices
- Inadequate GPU compute resources
- Inadequate computing power in general
- Require secure storage and analytic environments, allow external collaborators access. E.g. UK SeRP operated by Swansea University
- Require encrypted storage system with granular permissions controlled by researchers
- Nectar space expires when the grant period is over. Project doesn't end when the funding ends.
- Pawsey is unreliable.
- Require server based video storage.
- Require audio and video collecting instruments with auto transcription service.
- Infrastructure to support big data, R Drive is not suitable for big data research.
- Required software not available from Curtin (software mentioned are: MPlus, E.F Palisade's Decision Tree, NVivo, LISREL, R)
- Lack support for off-campus/remote researchers
- Not enough computers in Humanities HDR Hub
- Specialised programming support would be nice

Next step

34 respondents indicated interest in further discussions on their ICT needs. The eResearch Special Interest Group will consider how best to follow up with the respondents.

Conclusion

The survey results suggested that current research infrastructure is meeting the needs of the majority of the respondents. However researchers need more data storage space, compute power, flexibility to access and manage research infrastructure.

An almost equal number of HDR students and staff responded to the survey and the results indicated that the needs of students and staff are largely the same. This indicates that it would be desirable to provide HDR students the same access to ICT systems and support as staff, including the ability to create and manage R Drive without supervisor oversight.

The types of research data collected by researchers have not changed much in the last 7 years, but more researchers are collecting data from multiple sources and involving international collaborations. ICT system and support need to recognise this.

A significant proportion of data are collected from computer simulation and sensors. Automated data capture and ingest into research data storage could simplify research workflow and enhance user experience.

The R Drive is being used by a majority of the respondents, but slow speed when accessing remotely and inability to be accessed by external collaborators are major drawbacks for researchers. Dropbox, Google Drive, and OneDrive, remain popular cloud storage solutions used by Curtin researchers.

19% of respondents store their research data in only one location. HDR students seem less likely to retain their data for the minimum recommended retention period of 7 years. Guidance and awareness raising of data management best practice could help rectify this.

Ability to retain data and keeping data secure are very important to researchers. Making data publicly discoverable remains a low priority despite pressure from external funders for publicly funded research to have publicly available research outputs.

Free text responses surfaced the need for support in using ICT to enable research. Many researchers are not aware of the support, resources and services available to them such as remote access to R Drive, advice on data retention, and data publication service.