Australian Research Data Commons

Research Software Capability in Australia

By *Dr Michelle Barker* and *Dr Markus Buchhorn*,
for the *Australian Research Data Commons*

December 2021

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative
The ARDC is enabled by NCRIS

DOI: 10.5281/zenodo.6335998

Suggested citation:

ARDC Ltd. (2021). Research Software Capability in Australia. Viewed online at: *http://doi.org/10.5281/zenodo.6335998*

CONTENTS

Cover Image — monsitj - 327680560 / bigstockphoto.com

EXECUTIVE SUMMARY

The Australian Research Data Commons (ARDC) provides the Australian research community and industry access to nationally significant, data intensive digital research infrastructure, platforms, skills and collections of high quality data. ARDC's strategic themes include a focus on research software, in recognition of its role as a critical enabler of research and a key driver of innovation and economic growth. To support this focus on research software, ARDC commissioned a survey in 2021 on the scale and distribution of software engineering or development capability within research organisations in Australia. This work has been conducted in close alignment with the strategic aims of the Research Software Alliance, as part of the shared vision that research software and those who develop and maintain it are recognised and valued as fundamental and vital to research worldwide.

This report analyses and contextualises the survey results against national and international studies to consider the extent to which this research software engineering capability may be meeting Australia's research needs. The 70 survey responses included representation from 18 universities; six National Collaborative Research Infrastructure Strategy (NCRIS) facilities, three national and/or state eResearch bodies; three research centres, such as research discipline consortiums, medical research institutes, archives, etc.; and two Publicly Funded Research Agencies (PFRAs).

The first section of the analysis focuses on quantifying the number of personnel to illuminate this critical part of research infrastructure. The survey results suggested that there are approximately 4,000 people (or 2,500 EFT) working in funded roles that provide for software development, engineering and maintenance, and approximately 2,000 people in unfunded roles, totalling 6,000 people. These figures should be considered to be conservative estimates, based on significant under-reporting. Comparison with a similar 2019 study shows that total numbers of staff and EFT have increased approximately fourfold. The numbers for unfunded staff are ten times higher than the 2019 results, a much larger proportional increase, and shows that unfunded positions now contribute at least a third of the total effort.

Using an estimate of the total Australian research workforce to be 100,000 people, the survey analysis suggests that there is 1 EFT per 40 researchers for software development, engineering and maintenance (in 2019 the same survey question resulted in a figure of 1 EFT per 100 researchers). 30% of respondents felt there was sufficient research software capability in the area that they were responding about, or close to it, whilst 56% did not. When this was further broken down, 60% of the responses with a disciplinary focus perceived capability to be slightly or severely inadequate, as did 75% of responses with a whole-of-university focus. 67% of respondents reported that they may recruit more of this capability, and 43% stated that they will recruit more personnel in the next one to three years.

The second part of the analysis examines issues related to guidance, exploring perceptions around sufficiency of skills to ascertain training needs and availability. The survey results show that 46% of respondents perceived that the skills of their research software capability were adequate. 78% of respondents answered "yes" or "maybe" to a question on whether these personnel had access to mechanisms to improve their skills. Analysis of the data on personnel available to provide this guidance suggests that there are approximately 1,200 people working in funded roles (or 850 EFT) providing advice and training on relevant tools and applications, and approximately 300 in unfunded roles, totalling 1,500 people. This comparison shows that total numbers of staff and EFT have increased approximately twofold, and unfunded staff have increased sixfold, in comparison to 2019 data, with unfunded trainers providing about 20% of the personnel providing advice and training.

EXECUTIVE SUMMARY

This equates to 1 EFT per 118 researchers providing advice and training on relevant tools and applications, as compared to 1 per 200 researchers in 2019. It should be noted that the survey questions on training capability related development of all levels of research software skills, not just the more specialised skills utilised by the 6,000 people engaged in software development, engineering and maintenance.

The third section of the analysis focuses on community building, which includes analysis of common jobs titles for these personnel. A key aspect of community formation is enabling a sense of belonging, and the use of common job titles can assist staff in identifying themselves as part of a community. 80 different job titles were listed in survey responses as commonly used for these staff identified with 39% of these being titles commonly utilised for traditional academic roles (lecturer, professor, etc.). Exploration of career paths for research software personnel also yielded the information that options are scarce in Australia research organisations.

The fourth section of the analysis examined the survey information relating to advocacy, including policy and strategy. Questions that focused on organisational policy, particularly with regard to recognition of staff skills and software assets, showed that in comparison to international exemplars Australian research organisations have limited policies at the institutional level on research software and the capability that supports this. Funding arrangements for research software personnel were also surveyed, finding that only 33% of these staff have permanent employment.

The results of this survey are useful in understanding the extent to which this research software engineering capability may be meeting Australia's research needs. The analysis indicates there is a sizable and growing community research software support capability that needs to be developed, retained, sufficiently skilled and valued. However, the responses to questions such as numbers of staff, sufficiency of staffing, and intention to recruit indicate the community is not as well-served as it could or should be. Further investigation would be beneficial to provide increased understanding of some areas.

Technology is always advancing, with advances such as quantum and exascale computing set to enable a higher level of performance in research computing that has the potential to rapidly advance research impact. It is crucial that the human capital element of research software infrastructure also achieves the step changes needed to ensure there are personnel to support this technology.

LIST OF ACRONYMS

ABS -
Australian Bureau of Statistics

ADACS -
Astronomy Data and Computing Services

ANZSRC -
Australian and New Zealand
Standard Research Classification

ARC -
Australian Research Council

ARDC -
Australian Research Data Commons

ATN -
Australian Technology Network

CoE -
Centre of Excellence

CRC -
Cooperative Research Centre

DESE -
Department of Education, Skills and Employment

DORA -
Declaration on Research Assessment

ESFRI -
European Strategy Forum
on Research Infrastructures

EFT -
Equivalent-full-time

EOSC -
European Open Science Cloud

FAIR -
Findable, Accessible, Interoperable Reusable

FoR -
Field of Research

Go8 -
Group of Eight

QCIF -
Queensland Cyber Infrastructure Foundation

IRU -
Innovative Research Universities

MRI -
Medical Research Institute

NCRIS -
National Collaborative
Research Infrastructure Strategy

NHMRC -
National Health and Medical Research Council

OECD -
Organisation for Economic
Co-operation and Development

RSE -
Research Software Engineering

RSEs -
Research Software Engineers

NCRIS -
National Collaborative Research
Infrastructure Strategy

OSPO -
Open Source Program Office

PFRA -
Publicly Funded Research Agency

RBG -
Research Block Grant

ReSA -
Research Software Alliance

RUN -
Regional Universities Network

UK -
United Kingdom

UKRI -
UK Research and Innovation

UNESCO -
United Nations Educational,
Scientific and Cultural Organisation

1. INTRODUCTION

The Australian Research Data Commons (ARDC) provides the Australian research community and industry access to nationally significant, data intensive digital research infrastructure, platforms, skills and collections of high quality data. ARDC's strategic themes include a focus on research software, in recognition of its role as a critical enabler of research and a key driver of innovation and economic growth. To support this focus on research software, ARDC commissioned a survey in 2021 on the scale and distribution of software engineering or development capability within research organisations in Australia. Information about Australia's research software capability is currently limited, as is the case for most countries, and this survey aimed to provide some of the information about the capability needed to achieve a step-change in research that can improve Australia's standard of living, strengthen economic standing and build sovereign capabilities to protect national interests. This work has been conducted in close alignment with the strategic aims of the Research Software Alliance, as part of the shared vision that research software and those who develop and maintain it are recognised and valued as fundamental and vital to research worldwide.

This report analyses and contextualises the survey results to consider the extent to which this research software engineering capability may be meeting Australia's research needs. Section 2 explains the context for this work. Section 3 details the methodology, including the processes used to gather and analyse survey data, and presents the demographics of the survey responses. Section 4 explains the framework used for the analysis, explores the results, and contextualises against other national and international studies. Section 5 concludes with discussion on potential areas for future work.

2. CONTEXT

ARDC's interest in the extent to which the scale and distribution of software engineering or development capability is meeting Australia's research needs reflects both national and international drivers. The international research sector is increasingly recognising the value of research software and the people who develop it and maintain it. The need for this focus has been repeatedly demonstrated. Most researchers state that their work would simply not be possible without the use of software (Hettrick et al., 2014; Nangia & Katz, 2017). Despite this, the people who develop and maintain research software are not well incentivised and rewarded, with career paths being a particular issue, and these skillets are often highly desired and better compensated in industry (Berente et al., 2017; EOSC, 2021; European Commission, 2017; Scroggins & Pasquetto, 2020).

As the recognition that research software is a crucial part of research grows internationally, a growing number of governments are creating strategies and programs to ensure that research software and its personnel are supported to maximise benefits for the economy, environment and society (Barker et al., 2021). Australia is also progressing along this path. In 2021 ARDC released a draft National Agenda for Research Software to support increased recognition of research software as a first-class research output, alongside journal papers and datasets, and for recognising the importance of software as a research enabler (Honeyman & Treloar, 2021). Progression of this agenda will assist in ensuring that Australia will be able to fully utilise the value that research software represents.

ARDC's emphasis on research software is reflected in emerging Australian government priorities, with the National Research Infrastructure Roadmap Exposure Draft released by the Department of Education, Skills and Employment (DESE) including a focus on research software as part of its aims to provide a roadmap and vision for the next five to ten years (DESE, 2021). The National Research Infrastructure Roadmap Exposure Draft recognises that research software plays an essential but often invisible and undervalued role in generating, processing and analysing data. It also identifies emphasis on human capital as an emerging trend, with both technical expertise and a skilled workforce becoming increasingly important.

3. METHODOLOGY

A range of factors were considered in the development of this survey to assess the scale and distribution of research software capability. This section describes the survey design and distribution, and demographic analysis of the survey responses. This includes a breakdown of the 70 responses by organisational type and research discipline. Explanation is then provided on how responses were scaled to enable comparison across responses that ranged from individuals to the whole of organisation in their scope.

3.1 Survey Design and Distribution

This survey aimed to provide ARDC with relevant information on the research software capability in Australia, ideally in a way that would facilitate comparison with other contexts. One important step in designing the survey was to identify similar surveys used elsewhere, to identify questions that could be re-used in this survey. This is a limited field and only a few questions from other surveys have been reused, with their data used for comparison in the analysis in the next section.

One of the biggest challenges was clearly articulating what was in scope for the survey, as there are challenges in identifying the people who develop and maintain research software. This survey aimed to obtain an accurate indication of the size of the national workforce involved in supporting research software, which is a capability that is currently spread across professional and academic roles. People in the research community who develop research software range from researchers who may not consider themselves software developers at all, to people who combine professional software expertise with an understanding of research. Figure 1 depicts the overlap between roles that may include skills in research software development, including researchers and Research Software Engineer (RSEs). RSEs tend to combine professional software engineering expertise with an understanding of research processes.

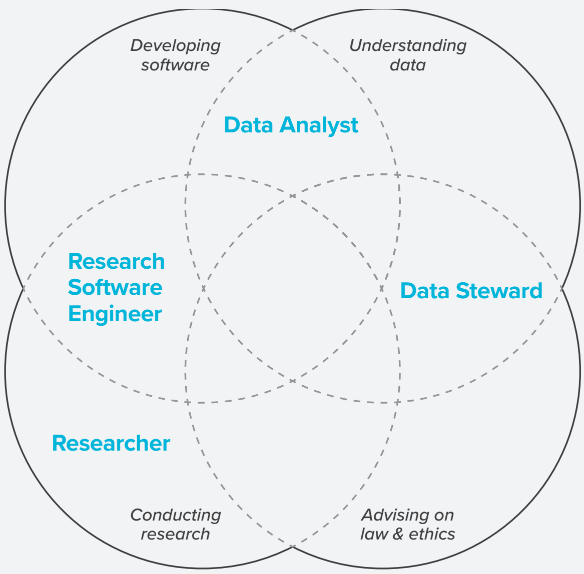

Figure 1: Digital workforce roles and responsibilities (OECD, 2020)

It is also difficult to gather information on the existence and distribution of different types of personnel with research software capability, as these personnel can be spread across many parts of an organisation and have many different job titles. The survey information explained that the focus was on staff with computing skills who support researchers in the development/engineering of software tools to facilitate the generation of research outcomes; and researchers/students who undertake such work, noting that this may not be a formal part of their job description. The roles of these staff could include coding, documentation, project management, software maintenance, but did not include research software trainers, business analysts, helpdesk staff, etc.

3. METHODOLOGY

The Australian research sector is broad and consideration was also given to which stakeholders the survey should target. The survey information stated that responses were sought from across the Australian research sector, including research institutions and supporting eInfrastructure providers, such as National Collaborative Research Infrastructure Strategy (NCRIS) facilities, universities, research centres and institutes, Publicly Funded Research Agencies (PFRAs), and regional/state/national eInfrastructure organisations. The survey asked for information about research software professionals at the respondent's institution, organisation, or group, as appropriate. Responses were encouraged from group/project leads, (senior) managers, chief/principal investigators or academic leadership, who were able to complete the survey on behalf of their unit or organisation as relevant, noting that in some cases it may be appropriate for individuals to respond in this regard.

The survey was open from 11 October to 19 November 2021, and distributed through emails, newsletters and social media. It contained questions about the size and capability of the area that the respondent was responding about, and about research software management and career progression pathways at their organisation. There were also spaces to provide additional information or clarification. A copy of the survey is contained in Appendix A.

3.2 Survey Responses

A total of 70 responses were analysed, some of which addressed the work of more than one organisation. The following organisations were represented, although not necessarily all at organisational level:

- 18 universities (providing 49 responses)
  - Responses from the 18 universities included:
    - 7 of the 8 universities in the Group of Eight (Go8)
    - 3 of the 8 universities in the Innovative Research Universities (IRU)
    - 2 of the 7 universities in the Regional Universities Network (RUN)
    - 2 of the 5 universities in the Australian Technology Network (ATN)
  - Of the 49 responses from universities:
    - 36 were from parts of the university that focused on a particular research discipline or focus, such as faculties or schools; research institutes, clusters, alliances, initiatives or groups (and included 3 responses from individuals)
    - 7 were estimates of the entire institution
    - 6 were from parts of the university that provided technology services across the institution, such as eResearch, technology or informatics centres/hubs/services
- 6 NCRIS facilities (providing 8 responses)
- 3 national and/or state eResearch bodies (providing 4 responses)
- 3 centres, such as research discipline consortiums, medical research institutes, archives, etc. (providing 4 responses)
- 2 PFRAs (providing 8 responses)

3. METHODOLOGY

The lack of responses from research centres is an obvious weakness in the spread of survey responses. Research centres such as Australian Research Council (ARC) Centres of Excellence (CoEs), Cooperative Research Centres (CRCs), Medical Research Institutes (MRIs) and National Health and Medical Research Council (NHMRC) Centres of Research Excellence, were poorly represented. For example, none of the 24 active CRCs at the time of writing were represented in the survey responses. Many of these organisations or centres sit outside universities, but some span multiple universities so may have been partially included in some other responses. For example, just one of the Go8 universities lists around 70 significant 'centres and institutes' that they host, plus nearly 20 centres supported through ARC and NHMRC programs, five CRC nodes and another seven centres funded through international bodies. Given the scarcity of whole-of-university responses to this survey, much of this information may not have been captured. It would be useful to augment the results of this study with further investigation of research software capability at research centres.

The areas for which respondents provided information ranged from entire organisations to small teams, and Figure 2 shows the number of researchers in that area, including postgraduate students. Respondents who answered zero were all from organisations or teams that provided services to researchers, but were not researchers themselves (with the exception of one response that was from an individual).

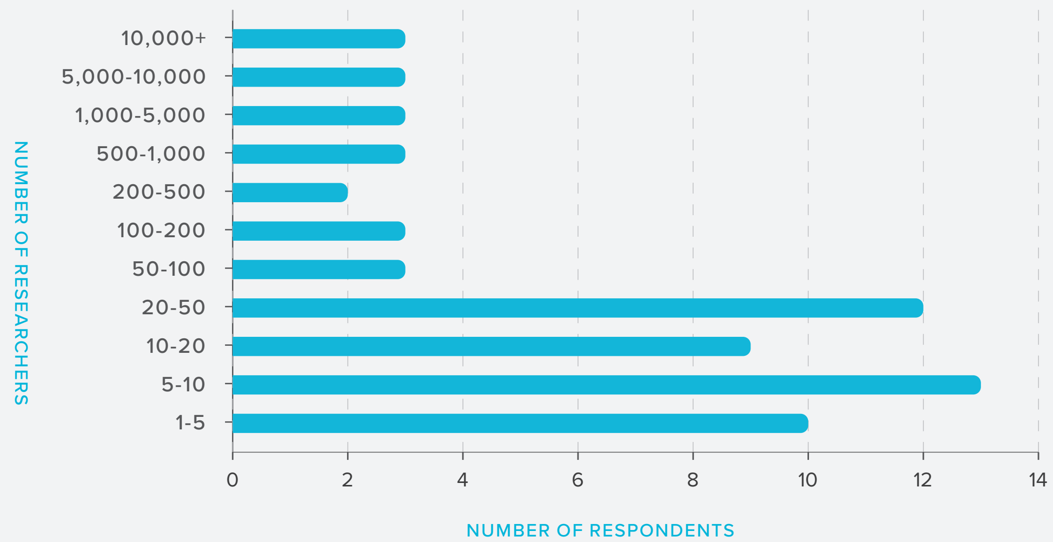

Figure 2: *Number of researchers in each respondent's area*

The 70 responses were also broken down into two groups based on the focus of the area:

- 50 responses were from areas with a disciplinary focus, i.e., an organisation, centre or team that focused on a particular research discipline or research focus.
- 20 responses were from areas with a whole of organisation focus, i.e., an organisation, centre or team that provided technology services to an organisation, or a whole of organisation response.

3. METHODOLOGY

This breakdown enabled differentiation of responses into the groups 1) the research software capability was embedded in a team or part of the organisation that was focused on a particular research area (which was often multi-disciplinary), and 2) the research software personnel were centralised and made available to researchers (who could be internal or external to the organisation) under a variety of service models (from merit based to fee-for-service). Universities, NCRIS facilities and PFRAs often utilise a combination of both models, whilst examples of the latter include state-based eResearch infrastructures such as Queensland Cyber Infrastructure Foundation (QCIF) and Intersect, who provide services to member universities in their states.

The survey responses were spread across a range of research disciplines, based on the two-digit Fields of Research (FoR) codes that comprise part of the Australian and New Zealand Standard Research Classification (ANZSRC), as shown in Figure 3[1].

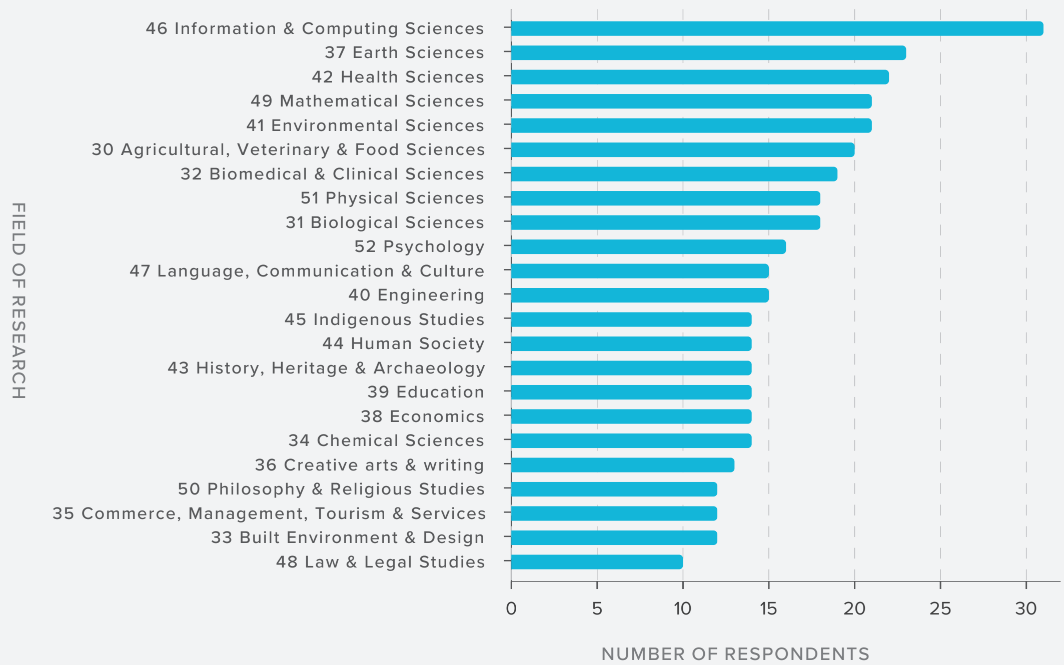

Figure 4: *Disciplinary-focused survey responses by FoR codes*

This shows that all disciplines were represented, with a number of FoR codes relating to science, technology, engineering and maths receiving higher representation. It is not surprising that the most commonly cited FoR code was informatics and computing sciences, as respondents sometimes worked in computing teams that served other areas of an organisation, or were part of research areas where computing was part of their multi-disciplinary team.

1. Three responses did not select any of FoR codes but wrote comments that they had a broad base, addressed all domains or incorporated almost all of the FoR codes. These responses were counted as addressing all FoR codes.

3. METHODOLOGY

This data was further broken down to differentiate between responses from organisations, centres or teams with either a disciplinary or whole of organisation focus. The breakdown for the 50 organisations with a disciplinary focus are shown in Figure 4, as this removes the responses where all (or almost all) FoR codes were selected.

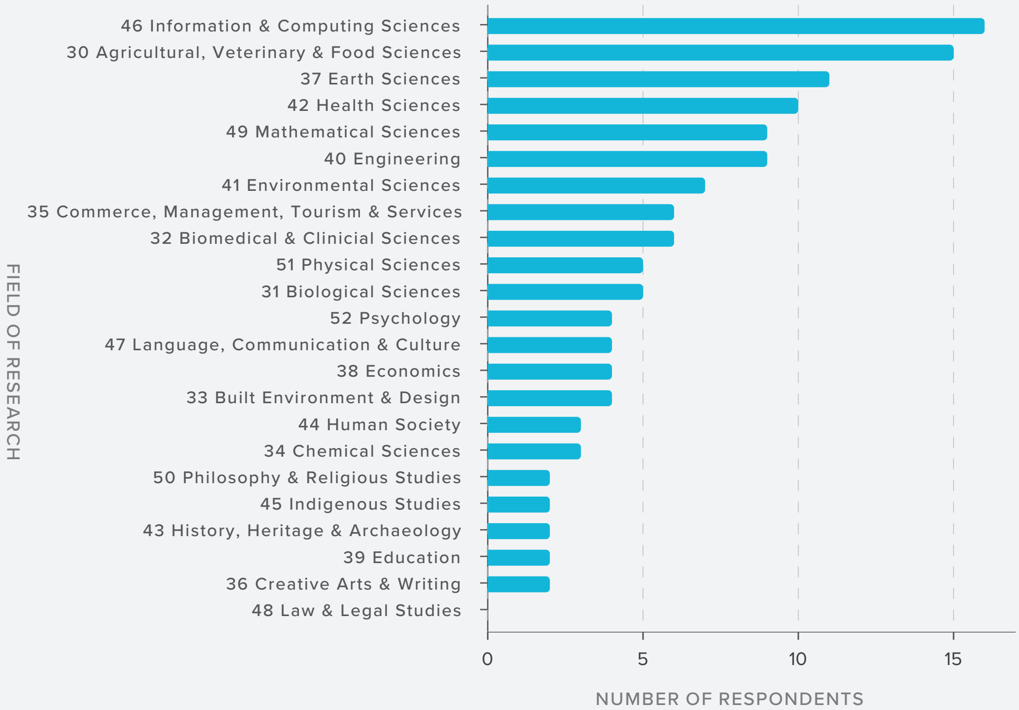

Figure 4: *Disciplinary-focused survey responses by FoR codes*

In this analysis it becomes more obvious that some FoR codes were better represented in the survey responses, such as agricultural, veterinary and foods sciences. The least well represented research areas were law and legal studies (zero responses); followed by creative arts and writing; education; history heritage and archaeology; indigenous studies; and philosophy and religious studies (one response each).

The humanities and social sciences are often less well represented in studies about research software use. The 2018 Australian RSE survey results detailed the disciplines in which the RSE respondents work, with 5% of less respondents nominating law and legal studies; commerce; economics; built environment and design; philosophy and religious studies; history and archaeology; studies in creating writing and writing; psychology and cognitive sciences; studies in human society; language; and education. In contrast, science, technology, engineering and maths disciplines scored well: 45% nominated information and computing sciences; 36% physical sciences; 34% biological sciences, etc. (May et al., 2019).

3. METHODOLOGY

Similarly, an international study found that whilst 85% of social scientists use software, only 10% have developed their own (Duca, 2019). These results contrast starkly with a United Kingdom (UK) study across disciplines that found that 92% of researchers use software and that 56% develop their own (Hettrick, 2014). An OECD report provides a more detailed view with an international analysis of the amount of scientific production that results in new data or code by disciplines (Bello & Galindo-Rueda, 2020, figure 3.4), as shown in Figure 5.

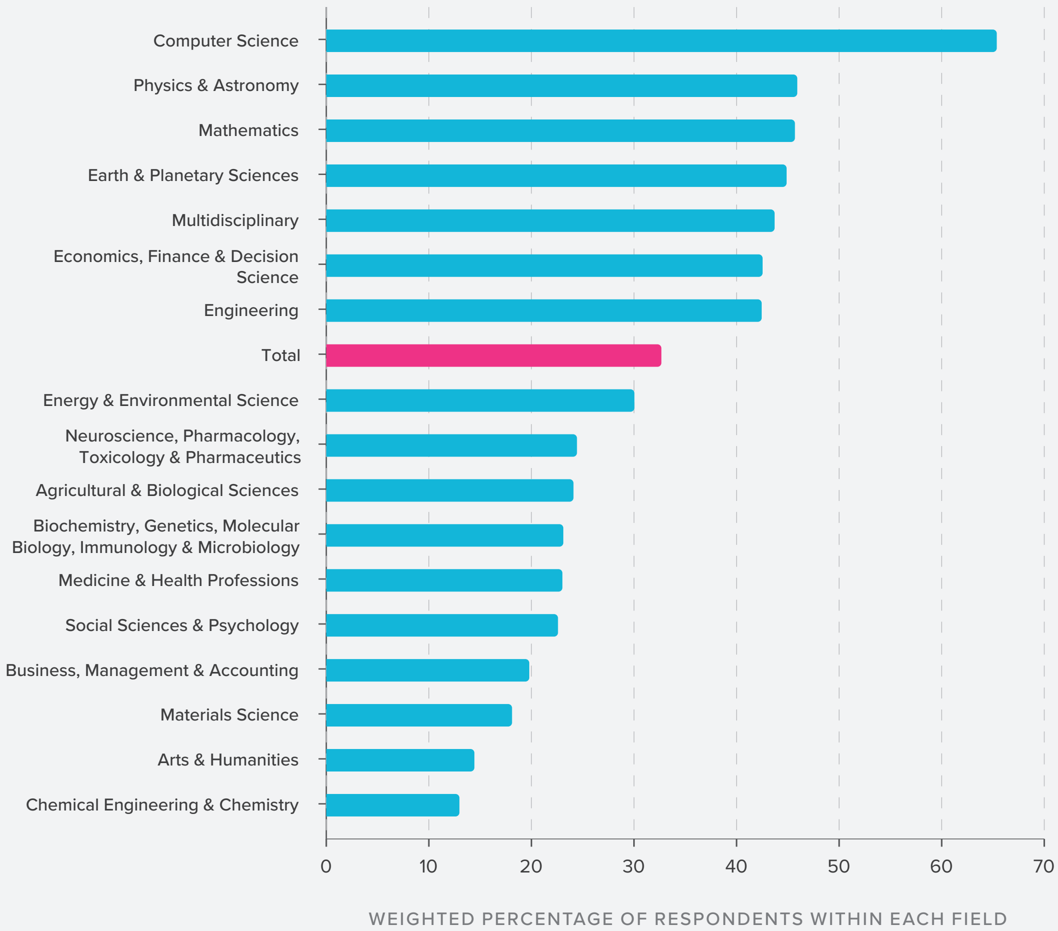

Figure 5: *Discipline breakdown of scientific production resulting in new code*

This aggregation of national results shows that in a number of disciplines (particularly computer science) more than 40% of the research produces code, while the figure is below 20% in other areas.

3.3 Scaling of results

The survey results were scaled to provide meaningful totals for the Australian research sector as whole. This was a complex process as the survey respondents comprised only part of the sector, were spread across a diverse range of organisations, and in most cases were partial rather than whole of organisation responses.

The process to extrapolate to likely totals across the country had to ascertain what scaling figures were reasonable, based on their category of organisation as per the five categories listed below. As a few organisations fit more than one category, in this analysis the organisations were clustered by their characteristics, e.g., if they were disciplinary focused, or had a whole of organisation focus.

1. NCRIS facilities can largely be enumerated, but their diversity in scale is significant. Some are major operational centres with a larger number of staff, while others act more as coordinating bodies for a range of institutions and centres. The responses represented approximately a fifth of all the NCRIS groups, and mostly from NCRIS facilities that were major operation centres, resulting in application of a multiplier of 5. It should be noted though that NCRIS responses contribute less than 3% of the staff count totals for this survey, so the impact on the totals of scaling is quite limited.

2. National and/or state eResearch bodies can also be enumerated, so their proportion of responses across the sector were calculated. The responses received represent approximately two-thirds of this sector and span the diversity of functions, consequently a multiplier of 1.3 has been applied.

3. Research centres, such as research discipline consortiums, medical research institutes, archives, etc. can only be counted as reported, with no scaling easily derived without a national catalogue of such bodies. These organisations contributed less than 2% of the reported totals, and this is certainly a gross under-count of their contribution across the sector.

4. PFRAs can be enumerated. The responses here covered about a quarter of the research workforce in this sector (based on their latest annual reports of staff counts), so assuming a reasonably consistent ratio of support staff to researchers, a multiplier of 4 has been applied.

5. Universities were a category for which the analysis was complex. There were three response types to deal with:

   - Whole-of-institution: Where no other data was provided from elsewhere within the institution, the response was taken at face value, with noted significant uncertainties and almost certain under-counting.
   - Part(s) of institutions: Based on experience with most of the responding institutions it was estimated that these centres reflected anywhere from a tenth to a third of their entire institution. For this analysis an average multiplier of 4.5 was used (with some larger institutions likely balancing out much smaller ones), noting that the contribution of these responses to the totals was usually smaller than the whole-institution responses. This multiplier is still likely to have generated an under-estimate.
   - A combination: In some cases a combination of institution-wide estimates were received, in addition to response from some centres and functional areas within the institution. These were checked for compatibility and/or overlap, and were then combined for their totals.

3. METHODOLOGY

Finally, to extrapolate totals across all universities when only a subset of universities were represented in the survey responses, an approach to correlating investment in research support with institutional research investment/income was used. This involved assumptions that could be overcome with more data, but provided a useful starting point for scaling.

DESE publishes a range of research intensity metrics annually, which in turn are aggregated to determine annual Research Block Grant (RBG) allocations (DESE, 2020). The universities that responded to this survey covered 75.3% of the national RBG allocations, indicating that a multiplier of 1.33 would be a reasonable extrapolation. It's also probably very conservative extrapolation, given the shortage of responses from institutional centres, and the rough estimates provided for whole-of-institution responses. The appropriateness of this multiplier was also checked by dividing the survey responses into whole-of-university responses (23.7% of RBG, a multiplier of 4.22) and partial university responses (51.6% of RBG, a multiplier of 1.94). Splitting the partial and whole institutions produced a reasonably consistent result, better than the simplest total of all responses and scaling. In this report the two results from the split approach have been averaged.

4. ANALYSIS

This consideration of the extent to which this research software engineering capability may be meeting Australia's research needs is framed in terms of the four levels used in the ARDC research software agenda of infrastructure, guidance, community and advocacy (Honeyman & Treloar, 2021). The scope and relevance of each of these is as follows:

1. **Infrastructure** includes the systems to store, publish, preserve, discover, develop, and (re) execute software, the software itself as a form of intangible asset, and the people (as soft-infrastructure) tasked with developing and maintaining software for broader use. This infrastructure will make change possible.
2. **Guidance** includes materials, training materials, reports and evidence gathering, presented as consultancy, support, events, webinars, workshops, and training. This will make adopting changes easier.
3. **Community** refers to software handling (concerned with publishing, access, discovery, citation, preservation) support communities, and software authoring communities (concerned with development and engineering concerns). Community formation leads to the shared behaviours and development and utilisation of shared assets that will make the change normative.
4. **Advocacy** includes funder and research institution policy and strategic vision, capturing the ways in which research software is required or incentivised, to create change in policy and strategy. This will initially create an environment conducive to change and eventually make the change codified.

This section uses these four levels to structure analysis of the survey responses, and contextualises the findings with national and/or international studies where possible.

4.1 Infrastructure

Research software infrastructure includes both digital infrastructure and people; the Australian government defines national research infrastructure as comprising "a range of nationally significant assets, facilities and services that support leading-edge research and innovation. However, it is much more than instruments and devices: a highly skilled workforce supports both the equipment and the researchers that use it." (DESE, 2021) This survey focused on the human capital elements of infrastructure, and this subsection explores quantification of Australia's research software capability, the sufficiency of this capability, and whether survey respondents plan to extend their staffing in this area. This information will assist ARDC in implementation of research software agenda recommended actions such as "ensure enduring research software infrastructure is available to researchers through a skilled, diverse, well distributed and sustainable workforce of research software engineers" (Honeyman & Treloar, 2021).

4.1.1 Quantification of capability

Survey respondents were asked to total how many people develop, engineer and/or maintain software used in research analysis to provide the data needed to approximately quantify the research software capability in Australia. The survey requested inclusion of people who primarily work with researchers, such as employees, contractors, interns, researchers and students. Respondents were also asked to answer in terms of numbers of people, equivalent full-time (EFT) staff, and/or unfunded people. The unfunded count aimed to identify people who are doing this work but for whom it is not their main responsibility (such as researchers and students). The raw totals were then scaled (as explained in the previous section) to gain estimates for the Australian research sector as a whole.

4. ANALYSIS

Table 1 summarises the survey results, showing both the raw data straight out of the survey, and the scaled figures used for the estimated national totals. The survey used a range selection to make it easier for respondents to report; to simplify the calculations here the midpoint of the range was used. Table 1 also breaks down this data across the five categories of organisations identified in the survey, i.e., universities, NCRIS facilities, national and/or state eResearch bodies, research centres, and PFRAs.

	PEOPLE (raw)	EFT (raw)	UNFUNDED (raw)	PEOPLE (scaled)	EFT (scaled)	UNFUNDED (scaled)
UNIVERSITIES	1276	747	880	3370	1957	2060
NCRIS	24	24	6	120	120	30
eRESEARCH	32	32	6	40	40	10
CENTRES	14	9	6	14	9	6
PFRAs	70	52	22	350	260	110
TOTALS	1416	864	920	3894	2386	2216

Table 1: *Counts of people who develop/engineer/maintain software used in research analysis*

Table 1 suggests that there are approximately 4,000 people working in funded roles (or 2,500 EFT), and 2000 in unfunded roles, totalling 6,000 people. These figures should be considered to be conservative estimates, based on significant under-reporting. Other mechanisms for calculating Australia's research software capability would suggest much higher numbers. For example, 56% of UK researchers were found to be developing their own software. Whilst the ARDC survey was narrower in focusing on personnel involved in tools creation and maintenance, the UK figures may still suggest under-reporting in the ARDC results. Similarly, OECD findings indicate that nearly 33% of research produces new code (Bello & Galindo-Rueda, 2020), and include an Australian figure of nearly 25% (OECD, 2019, section 5.2) . Any under-reporting may reflect a range of factors, including the difficulties in identifying the breadth of staff with this capability, and the lack of survey responses from research centres.

These indications are borne out from experience in the sector. Some projects in Australia in disciplines such as astronomy, environmental and biosciences, have made it clear that anywhere from a third upwards of the research staff actively develop a range of software, writing new tools or enhancing existing tools for their own requirements. In the information and computing science discipline it can be closer to 100%. For example, an analysis of Australia bioscience infrastructure elements which categorised researchers in terms of their data and technique intensity, found that significant changes were expected to occur over the next five years as more bioscience researchers required data-intensive and bioinformatics-intensive skills (Nisbet & Gray, 2018).

4. ANALYSIS

Table 2 provides the 2019 results for the same question (Buchhorn, 2019) and compares the totals with that for this 2021 survey. The 2019 survey received 42 responses from 18 universities, nine research centres (which included NCRIS facilities), and four national and/or state eResearch bodies.

	PEOPLE *(raw)*	EFT *(raw)*	UNFUNDED *(raw)*	PEOPLE *(scaled)*	EFT *(scaled)*	UNFUNDED *(scaled)*
UNIVERSITIES	268.5	239.5	124.5	430 - 860	380 - 760	200 - 400
eRESEARCH	45.5	45.5	0	90	90	0
CENTRES	48	48	8	50	50	10
TOTALS	362	333	132.5	570 - 1000	520 - 900	210 - 410
2021 TOTALS	1416	864	920	3894	2386	2216

Table 2: *2019 survey results with 2021 totals included*

This comparison shows that total numbers of staff and EFT have increased approximately at least fourfold since 2019. This is not surprising, given the increasing recognition of the importance of research software in research worldwide (Barker et al., 2020).

The numbers for unfunded people are ten times higher than the 2019 results, a much larger proportional increase. The finding that unfunded positions contribute at least a third of the total effort discovered here is not unexpected, with work in this field repeatedly identifying the significant role of volunteers in developing digital infrastructure (Eghbal, 2016). As noted by one of the survey respondents, open source research software capability can also exist beyond people associated with research organisations (even if unfunded), and that this survey also does not account for that labour. Open source software is software with source code that anyone can inspect, modify, and enhance. It is often maintained by a community, and in the case of open source research software, community members can easily extend to citizen scientists, unpaid members of the research community, software developers who work in industry or beyond.

It can also be seen from the 2021 survey data comparing EFT to people ratios that NCRIS facilities, and national and/or state eResearch bodies, show a higher proportion (essentially 1:1 people:EFT) of staff in full-time roles providing research software capability, while universities, research centres and PFRAs have proportionally more part-time workers (0.5-0.75:1) and unfunded workers (0.3-0.5:1). The whole-of-university survey responses also reported a significantly higher proportion of unfunded staff than the responses from parts of an institution, by a factor of around three, perhaps indicating that the former have greater visibility of unfunded staff.

4.1.2 Is the research workforce well-supported?

It is useful to compare the quantification of research software capability to the scale of the research community. The national higher education research and development workforce (excluding students) was estimated by the Australian Bureau of Statistics (ABS) at around 82,000 EFT in 2018. This includes post-graduate students (56% or 46,000 people), academics (30%) and 'other supporting staff' (14%) (ABS, 2020a); and government and non-profit data (who tend to collaborate with universities) which combines researchers with technical staff and 'other staff' (ABS, 2020b). DESE data identified 66,000 students studying higher degrees by research in 2019, which is nearly 50% more than the ABS data for postgraduates overall, so these have been averaged to increase the 82,0000 EFT figure to 92,000. Government resources devoted to research and experimental development include another 7,500 EFT of researchers (ABS, 2020b).

This provides an estimate of the total Australian research workforce as 100,000 people, and when combined with the survey data this suggests that there is 1 EFT per 40 researchers for software development, engineering and maintenance. In 2019 the same survey question resulted in a figure of 1 EFT per 100 researchers (Buchhorn, 2019), showing that numbers have more than doubled in 2021. Research providing quantification of research software capability is rare, partially reflecting the challenges in defining clear roles. One of the only other analyses available of digitally skilled research personnel is an analysis of data librarians at American R1 universities (doctoral universities with very high research activity) which found that the average number of data librarians per R1 university is a little over two (Springer, 2019). Whilst this is not easily compared with the results here, it does show the general lack of staff with digital research skills.

4.1.3 Sufficiency of capability

To determine the adequacy of existing research software capacity, survey respondents were asked how sufficient they felt existing capability in their area to be. Figure 6 shows the percentage of respondents answering across a range of sufficiency levels.

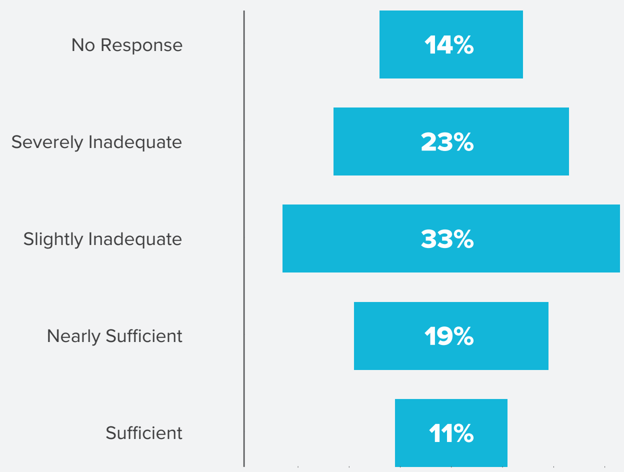

Figure 6: *Responses to sufficiency of this number to meet organisational needs*

Figure 6 shows that whilst 30% of respondents felt there was sufficient capability, or close to it, 56% did not. While it is difficult to understand the numbers of existing personnel with research software skills, it is even harder to estimate what level of staffing with this capability is needed, particularly as there are many different roles that can incorporate research software skills. The answers can also vary across research disciplines. For example, most researchers in bioinformatics acquire some research software skills in their undergraduate degrees, whilst this appears to be less common in some disciplines.

Disciplinary differences can also complicate consideration of the number of research software professionals needed. It has been suggested that 1 in 20 staff in the research workforce should be digitally skilled research support professionals (Mons, 2020). In Europe this translates to approximately 500,000 professionals of various kinds to support researchers through experimental design and data capture, curation, storage, analytics, publication and reuse.

Figure 7 further breaks down the preceding figures, differentiating by responses from a disciplinary or whole of organisation focus.

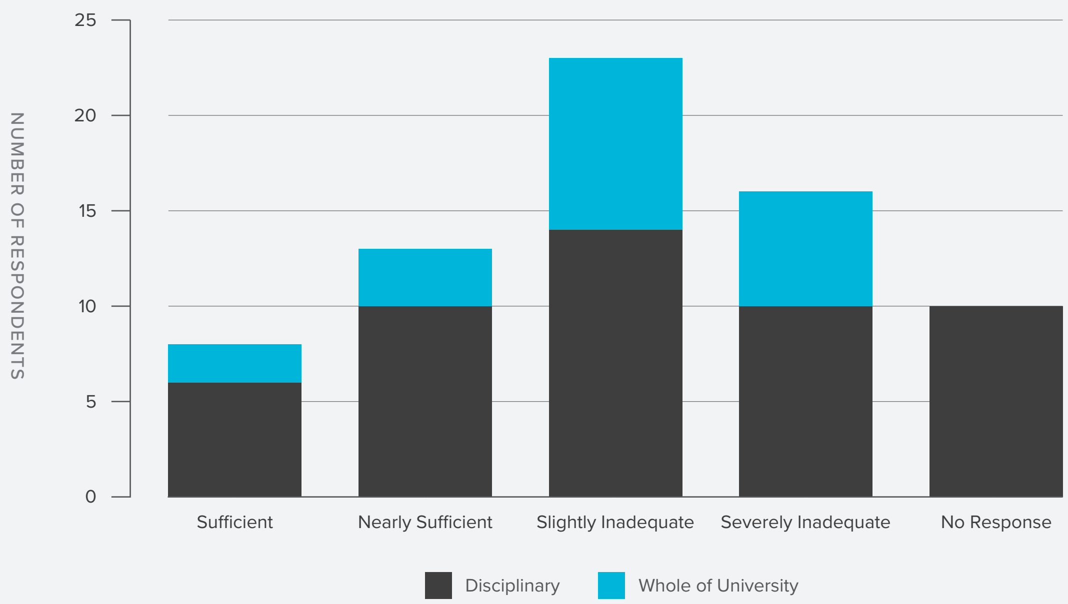

Figure 7: *Comparison of disciplinary or whole of organisation focus*

Figure 7 shows that 60% of the responses with a disciplinary focus perceived capability to be slightly or severely inadequate, as did 75% of responses with a whole-of-university focus. It should be remembered that the whole-of-university focus includes parts of the university that provided technology services across the institution, such as eResearch, technology or informatics centres/hubs/services. It is possible that the slightly higher perception of insufficiency of capability from responses with a whole-of-university focus is due to their bird's eye view of demand across much or all of the organisation.

This can be further refined to focus on those responses with a disciplinary focus, across research domains. In Figure 8 the responses for each FoR code are aggregated into five of the domains used by the European Strategy Forum on Research Infrastructures (ESFRI): environment, health and food, physical sciences and engineering, social and cultural innovation, and digital. This clustering has limitations, but after examination of a range of aggregations models was considered the best option as it potentially offered opportunity for comparison internationally.

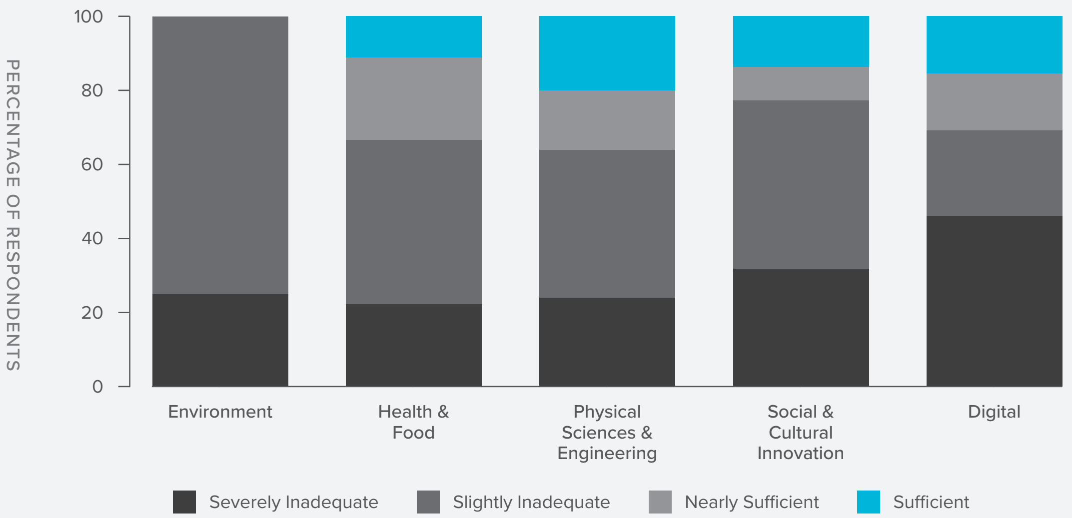

Figure 8: *Sufficiency by research domains*

Figure 8 suggests that responses that identified the digital domain (which corresponds to FoR code 46: Information and computing sciences) as part of all of their focus are most severely underserved (based on 14 responses). The environment domain aggregates only four responses, which was by far the smallest aggregation, so may not be statistically significant. If this is discounted, then the social and cultural innovation domain records the highest level of inadequacy overall, whilst physical sciences and engineering experience the most sufficiency.

4.1.4 Plans to increase capability

The next survey question asked whether respondents had plans to recruit more people with this capability in the short-medium term (one to three years), with percentage responses provided in Figure 9.

Figure 9 shows that 67% of respondents may recruit more of this capability, and that 43% will. This finding aligns with a range of reports that assert that digitals skills will be increasingly needed in the research sector, including the National Research Infrastructure Roadmap Exposure Draft. Mast provides another rationale for increased growth through measurement of the growth in Australia's eResearch activities, concluding that:

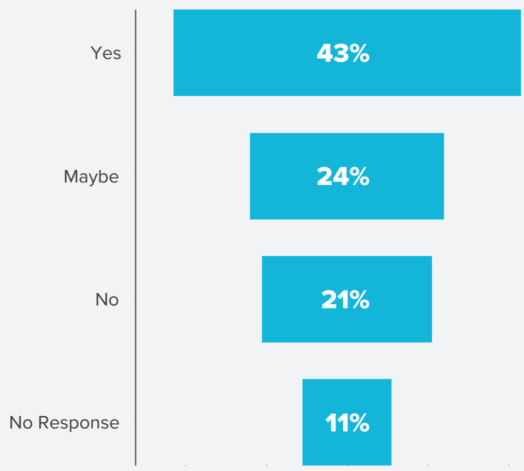

Figure 9: Plans to recruit in the short-medium term

The growth of data-related publications, expressed as a percentage of total publications, has risen from 11.7% to 22.1% globally. In Australia, the growth is considerable from 13.4% in 1989 to 28.0% in 2018, surpassing other R&D intensive countries such as the US, Canada, China and Singapore. Taking this trend as a baseline, it suggests that growth in data-related publications, and thereby eResearch activities, will likely continue in the near future (ARDC, 2019).

4.2 Guidance

Guidance such as materials, training materials, reports and evidence gathering, are needed to make it easier for both the research sector and research software capability to better recognise and value this human capital. To explore Australia's readiness to provide this to the research sector, several survey questions explored perceptions around sufficiency of skills, to ascertain training needs and availability. The results of these can help inform actions recommended in the ARDC research software agenda in this area, such as "connect researchers and research software engineers to guidance to develop research software infrastructure for easiest and broadest reuse" (Honeyman & Treloar, 2021).

4.2.1 Sufficiency of skills

Research on the software development training that people who develop research software require often shows that more skills are needed. This survey asked respondents to consider if their research software staff have the skills they need, with responses shown in Figure 10.

Figure 10 shows that 46% of respondents perceived that skills were adequate, with only 5% outright disagreeing with this. It may be relevant to remember that this survey question asked participants to consider the skills needs of their team, centre or organisation; rather than their own skills. In comparison, in the Stanford software survey where individuals assessed their own needs, 75% of participants felt that they had not received sufficient training for software engineering best practice. However, the majority did consider themselves to be at least proficient in terms of their software development expertise (Stanford Software Survey, 2020). It would be useful to understand in more detail the skills that are lacking, including capability across the four high level software engineering areas for skills development that the ARDC research software agenda identifies: encapsulation techniques, abstraction, performance, and integrity (Honeyman & Treloar, 2021). Both technical and soft skills are usually considered important, and the ARDC research software agenda also highlights both. The Australian data of the 2018 RSE survey identifies the top skills that participants wanted to improve as being project management, software testing, software design, machine learning, and Kubernetes (May et al., 2019).

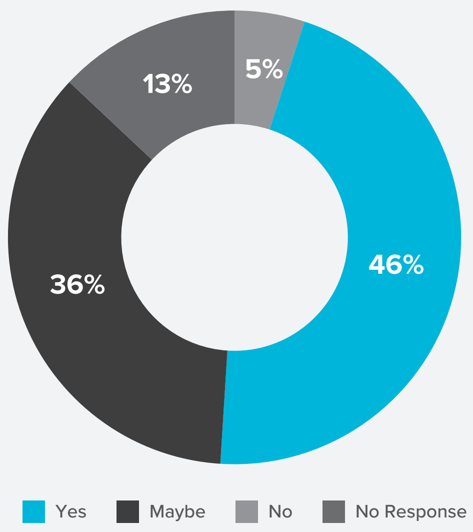

Figure 10: *Sufficiency of research software staff skills*

4.2.2 Sufficiency of training opportunities

A common area of concern is if there are adequate training opportunities for people who develop and maintain research software. This reflects that many staff come from research rather than computational science backgrounds, and may not have received much (or any) formal training in software development. At the other end of the spectrum, research software professionals must keep up with rapidly developing technology. To shed light on the Australian situation, survey respondents were asked if their research software staff have the mechanisms to acquire the skills they need, with answers shown in Figure 11.

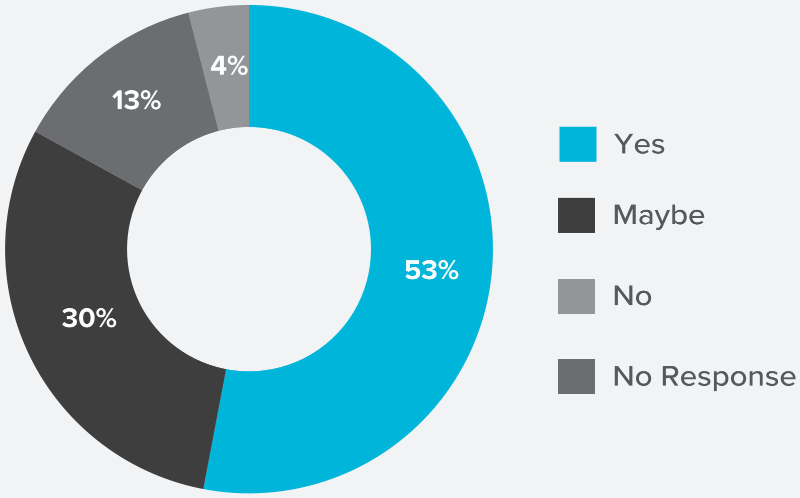

Figure 11: *Do your research software staff have mechanisms to acquire the skills they need?*

Figure 11 shows that 78% of respondents answered “yes” or “maybe”. Information provided in response to an open question on how research staff access training in this area yielded information included about range of options, including:

- Self-directed learning
- Mentorship, usually informal
- Communities of practice
- Internal training on standard topics, e.g., python programming
- Online courses from reputable and recognised institutions and course providers, including university micro-credentials
- Organisational general professional development e.g., leadership, communication courses
- Specialist courses offered by software/hardware vendors
- Attendance at industry events
- National and international collaboration

Participants named some useful training opportunities, including courses and/or communities of practice provided by the Carpentries, Python Charmers, national and state eResearch bodies, university eResearch centres, and disciplinary efforts such as the Astronomy Data and Computing Services (ADACS), who provide astronomy-specific training to allow astronomers to maximise the scientific return from data and computing infrastructure. Some of the examples given are initiatives that provide introductory skills in software development, such as the Carpentries, rather than the more specialised skills utilised by personnel engaged in software development, engineering and maintenance.

Whilst there was a positive response to the questions on both the adequacy of skill sets and availability of training, a number of open-ended responses emphasised the lack of structure around this. 9 of the 56 comments received emphasised that there was no formal training, and that individuals self-directed to seek training themselves. Comments about the difficulty in accessing useful training included the following:

- It may be difficult for some researchers to know what training would be helpful/effective for their staff.
- New hires are not made aware of training resources because line managers and leaders are not aware of training resources and don't have an understanding of software engineering.
- Training occurs on the job, during development. It's often difficult to justify the time spent on training in terms of outcomes which will further attract funding.
- It is a bit hard as different people/groups have different requirements, often software specific.

4.2.3 Quantification of capability to provide training

The research sector's ability to provide training in research software development is dependent to some degree on personnel available to provide this guidance, noting that some of the training comes from industry or other external providers. One survey question repeated another question from the 2019 survey, asking how many people provide advice, education/training to researchers/research teams on relevant tools and applications. It should be emphasised that this question assessed the personnel providing training at any level of research software development. Results are shown in Table 3.

	PEOPLE *(raw)*	EFT *(raw)*	UNFUNDED *(raw)*	PEOPLE *(scaled)*	EFT *(scaled)*	UNFUNDED *(scaled)*
UNIVERSITIES	300	191	89	1003	644	255
NCRIS	14	14	6	70	70	30
eRI	44	44	6	57	57	8
CENTRES	9	9	3	9	9	3
PFRAS	35	17	6	140	68	24
TOTALS	402	275	110	1279	848	320

Table 3: *Number of people providing advice, education/training to researchers/research teams on relevant tools and applications*

4. ANALYSIS

This suggests that there are approximately 1,200 people working in funded roles (or 850 EFT), and 300 in unfunded roles, totalling 1,500 people. Using the estimate of the total Australian research workforce of 100,000 people as described earlier, this analysis suggests that there is 1 EFT per 118 researchers providing advice and training on relevant tools and applications, as compared to 1 per 200 researchers in 2019. Table 4 compares this to 2019 data (Buchhorn, 2019).

	PEOPLE *(raw)*	EFT *(raw)*	UNFUNDED *(raw)*	PEOPLE *(scaled)*	EFT *(scaled)*	UNFUNDED *(scaled)*
UNIVERSITIES	187.5	121.5	21.5	300 - 600	200 - 400	30 - 60
eRESEARCH	55	42.5	0	110	80	0
CENTRES	21.5	13.5	4.5	20	10	5
TOTALS	264	177.5	26	430 - 730	290 - 490	35 - 65
2021 TOTALS	402	275	110	1279	848	320

Table 4: *2019 survey results with 2021 totals included*

This comparison shows that total numbers of staff and EFT have increased approximately twofold, and unfunded staff have increased sixfold. There has been a significant increase in the emphasis on training in digital skills for researchers in recent years, with open science providing a key driver. For example, one of the three main EOSC objectives is: “Ensure that Open Science practices and skills are rewarded and taught, becoming the ‘new normal’” (EOSC Executive Board, 2020). This has long been needed; an analysis of United States National Science Foundation principal investigators in biological sciences identified that the most unmet need was training (Barone et al., 2017). In 2020 the OECD identified issues in the extent to which available digital skills training fails to meet supply, and the difficulties for organisations in scaling up to meet the ever-increasing demand (OECD, 2020). The National Research Infrastructure Roadmap Exposure Draft notes: ”that some research areas and institutions are already building data and computational expertise internally as an essential resource. There should be consideration of how national system-wide approaches to training and services could benefit researchers.” (DESE, 2021)

The 2021 data shows that unfunded trainers provide about 20% of the personnel providing advice and training. This percentage is lower than expected, perhaps suggesting that there is still a lack of recognition of the contributions from postgraduate students and early career researchers. It is commonly recognised that the people providing training are usually not certified trainers, and often do not receive recognition for their expertise and contributions in training (OECD, 2020). In recognition of this, organisations such as the Carpentries provide not only foundational coding and data science skills, but also instructor training. At least 269 Carpentries courses have been convened in Australia since 2013 (The Carpentries, 2021); however, this data is likely to be incomplete as it only reflects information that organisations self-report to the Carpentries.

4.3 Community

Communities are an important part of the skills development in the research sector, enabling diverse groups of people with common interests to come together to achieve change at various levels (OECD, 2020). There are a wide range of communities that support development of the research software capability, with a study by the Research Software Alliance (ReSA) showing that almost 90% of research software initiatives focused on personnel issues include community in their remit. The focus of these communities varies, and includes sharing of best practice in particular areas, such as software citation or development of RSE career paths; and provision of mutual support for people identifying as part of emerging communities, such as RSEs, or particular types of trainers, learners, and leaders (Barker, 2021).

Whilst this survey did not specifically aim to address aspects of community building, community building is an important part of the Australian research software landscape. Some of the information gained from this survey can assist ARDC in building a range of targeted communities for research software personnel, including the understanding gained around common jobs titles for these personnel, and recognition of their skill sets through career paths. For example, the ARDC research software agenda recommends actions to be taken to achieve recognition of research software as a first class output of research including “build and sustain local, regional and national communities of practitioners who bridge the gap between research and software engineering practices”, and “build and sustain a national community of professional research software engineers” (Honeyman & Treloar, 2021).

Diversity, equity and inclusivity are also important considerations for community building in the research sector. There is a significant lack of diversity in the research software community; international analysis in 2018 on research software engineer demographics found that 73-92% were male, and most commonly ranged in age from 25-44 (Philipe, 2018). UK RSE data revealed a community that was 14% women, 5% black, Asian and/or minority ethnic/mixed in ethnicity, and with 6% reporting a disability. These figures were more than 50% less than the figures for the UK workforce as a whole (Chue Hong et al., 2021). Understanding of these issues in the Australian research software workforce would be useful, and future investigation could include these elements.

4.3.1 Common job titles for research software personnel

A key aspect of community formation is enabling a sense of belonging, and identifying as part of a community is a first step. The success of the RSE movement has been partly due to the growth of community members who identify as RSEs. One survey question sought to understand job titles commonly used for the staff identified in this survey, to understand whether common job titles are being used, as this could assist those staff to identify as part of a community.

Responses on job titles commonly used for these staff identified in this survey provided 80 different job titles, as listed in Appendix B. This diversity aligns with a 2019 analysis of roles associated with the Australian eResearch domain which utilised 180 job titles in an exercise to categorise them by the primary service that the role performs, then aggregated them into capabilities (May, 2019).

Where clustering is applied to these job titles, 39% of the mentions of job titles are commonly utilised for traditional academic roles: Postdoctoral Researcher (or Associate), types of Lecturers and Professors. Similarly, at least 22% relate directly to software development, including Junior Developer, Junior Software Developer, Senior Software Developer, Software Developer, Research Engineer, Research

Software Developer, Research Software Engineer, Senior Software Developer, Senior Software Engineer, Software Engineer, and Software Engineer (Research), although others could possibly be included, such as Informatics Software Developer, Scientific Programmer, etc.

The most commonly mentioned job titles were:

- Research Assistant or Research Associate – 10 mentions
- Postdoctoral Researcher or Postdoctoral Research Associate – 9 mentions
- PhD student – 7 mentions
- Software Developer – 6 mentions
- Lecturer – 5 mentions

The most common job titles is also similar to the results of the Stanford software survey which lists 12 job titles (across 65 respondents), with the most common being PhD student (25%), Postdoctoral Researcher (14%), Professor (14%), Software or Systems Engineer (12%) (Stanford Software Survey, 2020).

4.3.2 Levels of research software roles

The survey also sought to understand the career paths available for research software capability. The research sector is undergoing significant advancements internationally in how research is incentivised, supporting recognition of research software skills and outputs. Governments, funders, institutions, publishers and others are introducing policies aimed at making research outputs more open and research practice more inclusive and collaborative. The inclusion of appropriate incentives in the reward and evaluation system for researchers is critical to achieving this. The G7 Research Compact highlights actions which include exploring incentives that foster recognition and reward collaboration to drive a culture of rapid sharing of knowledge, including software and code (G7, 2021). International initiatives such as DORA, the Leiden Manifesto, and the Sorbonne Declaration are helping to change culture around research evaluation and metrics.

In the case of research software, there is a need to recognise and support a range of roles, from researchers who do occasional coding, to software development professionals–and everything in between. As Knowles, Mateen and Yehudi argue, research software "needs to be funded, maintained and have viable career paths even if the researchers involved are writing more lines of computer code than lines in an academic manuscript." (Knowles et al., 2021). To help understand what roles currently exist the survey asked how many different levels of research software staff with different responsibilities existed (e.g., junior and senior roles, team leaders, supervisors, managers), with Figure 12 showing the percentage of respondents who completed this question (excluding the 20% non-responses), with answers ranging from one to five or more.

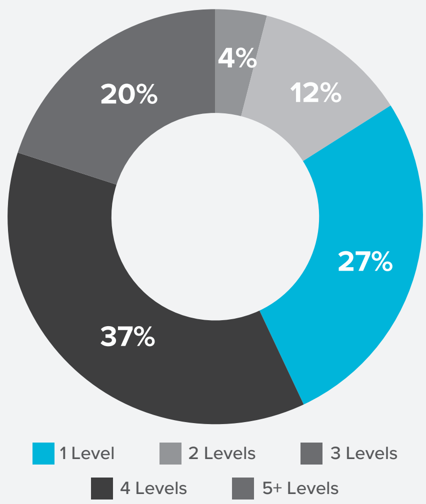

Figure12: *Number of levels of research software staff roles*

4. ANALYSIS

It is difficult to use this data to understand the potential for career path progression for these staff. While it is heartening that two or more roles were identified by 80% of those who responded to this question, some of these answers could be referring to the traditional academic career path discussed above, of postdoctoral researcher through to professor.

Whilst a number of respondents stated that there were a range of jobs with different responsibilities, it is challenging to identify different levels of job titles in the 80 job titles collected in this survey. Analysis of individual responses shows how roles can differ in terms of specialisation in specific organisations, for example:

- Data Manager, Data Management Specialist
- Informatics Software Developer, Software Developer
- Software Developer, Research Software Developer
- Software Engineer, Software Engineer (Research)

However, the range of titles in each organisation that reflect job specialisation are still limited. As one respondent noted: "Software engineering staff are given advanced research tasks that are not reflected in their job title or compensation level because managers do not understand what we do."

Analysis of individual responses also identified some organisations that may provide potential paths through some combination of junior, senior and manager roles, e.g.:

- Data Manager, Research Data Systems Manager
- Junior Developer, Software Developer
- Junior Software Developer, Software Developer
- PhD student, Associate Lecturer, Lecturer, Senior Lecturer, Associate Professor, Professor
- Software Engineer, Senior Software Engineer
- Software Developer, Software Engineer, Senior Software Developer, Senior Software Engineer

Career paths for staff with research software skills are being introduced in a range of institutions internationally. For example, the National Centre for Supercomputing Applications in the USA provides five levels in their career path for research programmers: Assistant Research Programmer, Research Programmer, Senior Research Programmer, Lead Research Programmer, and Principal Research Programmer. The University of Manchester RSE group has three levels of RSE, from junior to senior, that roughly equate to Research Assistant, Postdoctoral Research Associate and Lecturer on the university's academic track (Katz et al., 2019).

The National Research Infrastructure Roadmap Exposure Draft identifies as an emergent international trend "an increasing focus on digital research skills that deliver best practice in research data and software management. These essential staff need reward systems and career pathways that simultaneously support open science practice and their professional development." It includes a recommendation to develop an Expert National Research Infrastructure Advisory Group whose immediate priorities would be provision of advice on the development of a National Research Infrastructure Workforce Strategy to support career pathways, address technical skills shortages and identify capability gaps (DESE, 2021)

4.4 Advocacy

Advocacy refers to the authoring, implementation and advocacy for policy and strategy. Research institutions are identified as playing important roles in policy and strategy in the ARDC research software agenda, as employers of those that write or support the writing of research software, and as implementers of policy concerning software outputs, research institutions. It is highlighted that these organisations determine:

- How they support research software as an output
- How skills and capabilities are distributed, supported and nurtured
- How capabilities are crystalised in roles and structures in the organisation
- What strategic importance is given to research software as a research priority (Honeyman & Treloar, 2021).

To increase understanding of relevant policy in organisations responding to the survey, a number of questions focused on organisational policy at the highest level, and with regard to recognition of software assets. These are then contextualised with the findings with national and/or international studies where possible. This information will assist ARDC in implementation of research software agenda recommended actions including "create the policy and incentives environment that recognises research software as a first class output of research" and "Create the policy and incentives environment that supports the development and maintenance of critical research software infrastructure" (Honeyman & Treloar, 2021).

4.4.1 Organisational policy relating to research software

The ARDC research software agenda identifies that "a critical shift to enable change is the broadening of scope in policy to include software in its various forms as an output of research" (Honeyman & Treloar, 2021). Consequently, one of the survey questions asked: "What, if any, policy does your organisation have that relates to research software? (E.g., open science policy, software management plan policy, etc.)" None of the responses provided research software policies at the institutional level that are comparable with international exemplars. For examples, the Netherland's Delft University of Technology has a Research Software Policy that:

- Facilitates best-practices on research software management and sharing, irrespective of whether the code is proprietary or open source.
- Emphasises the value of research software as a standalone research output and facilitates proper recognition of the contribution of TU Delft researchers to software.
- Sets out some high-level requirements for how software should be managed, the responsibilities of the different stakeholders involved in software development and describes the global workflows that facilitate sharing software openly (Akhmerov et al., 2021).

36 survey responses did provide details of policies that included some elements related to research software, such as support for publication of software, encouragement of open source software, or that open science policies underpin their work. 25 respondents did not answer this question, and 16 respondents wrote comments that were not very encouraging, such as: "you have to stretch the rules to do this", "in general it is just done without asking", and "policy is unclear at best".

4. ANALYSIS

The importance of research software at the policy level is recognised by its inclusion in international standards such as the Organisation for Economic Co-operation and Development (OECD) revised Council Recommendation on Access to Research Data from Public Funding, and the United Nations Educational, Scientific and Cultural Organisation (UNESCO) Recommendation on Open Science. The recently revised OECD recommendation is a legal instrument that facilitates implementation by member states, and advocates for a range of practices to support research software, including training of research software engineers and enabling recognition and reward of software development skills (OECD, 2021). The UNESCO recommendation supports recognition of software and source code as part of open scientific knowledge, and encourages free access to open source software and source code (UNESCO, 2021).

A range of countries are now improving national policy initiatives and programs supporting research software, including Australia, Canada, European Union, France, Netherlands UK and USA.(Barker et al., 2021) An increasing number of countries and regions also have open science strategies that are driving integration of open science principles into research sector practices, including the European Commission, Canada, Japan, and some European countries (European Commission, 2019; Japan Science and Technology, 2017; Office of the Chief Science Advisor of Canada, 2020; SPARC Europe & Digital Curation Centre, 2019).

Another survey question sought to uncover whether organisations recognise software development or maintenance effort towards academic progression, as recognition of the contribution of research software developers to research outcome is one element needed to support this capability. 61% of survey respondents (from a total of 46 responses) answered that there was no way for this type of recognition to occur. However, five of the 28 responses stating this commented that this was because their research software staff were employed as professional or technical positions (or contractor roles). One respondent clarified that: “Software citations, publication acknowledgements, number of users of platforms are ways to report and recognise achievements, however this does not impact career progression.”

20% (or nine responses) identified that some sort of limited recognition was available, including:

- Informally.
- Limited to a particular team.
- As part of the wider research support recognition.
- Through publication of peer-reviewed manuscripts describing software.
- Promotion cases may quote impact form software, but much of our software is not shared, and it doesn’t count as publications.
- The promotions committee is prepared to listen to arguments around the contribution made by research software. How much weight is given will depend on the committee members at the time.

7% (or three responses) said their organisations did provide this recognition, although none provided details of an organisational wide policy. Examples were given that recognition was possible in some faculties (such as engineering or data science), or as a component of higher research degree completion. Only 20 Australian research organisations are signatories to the San Francisco Declaration on Research Assessment (DORA), which recognises the need to improve the ways in which researchers and the outputs of scholarly research are evaluated. Australian signatories include only one university (the University of Melbourne), Association of Australian Medical Research Institutes, Australian Academy of Science and National Health and Medical Research Council (DORA, 2021).

4.4.2 Employment types for research software personnel

The next set of questions sought to understand how these policies translated into recognition of the value of research software assets, in terms of both personnel and research software. As a starting point, respondents were asked the basis on which staff with this capability were employed, as shown in Figure 13 (excluding 16% non-responses).

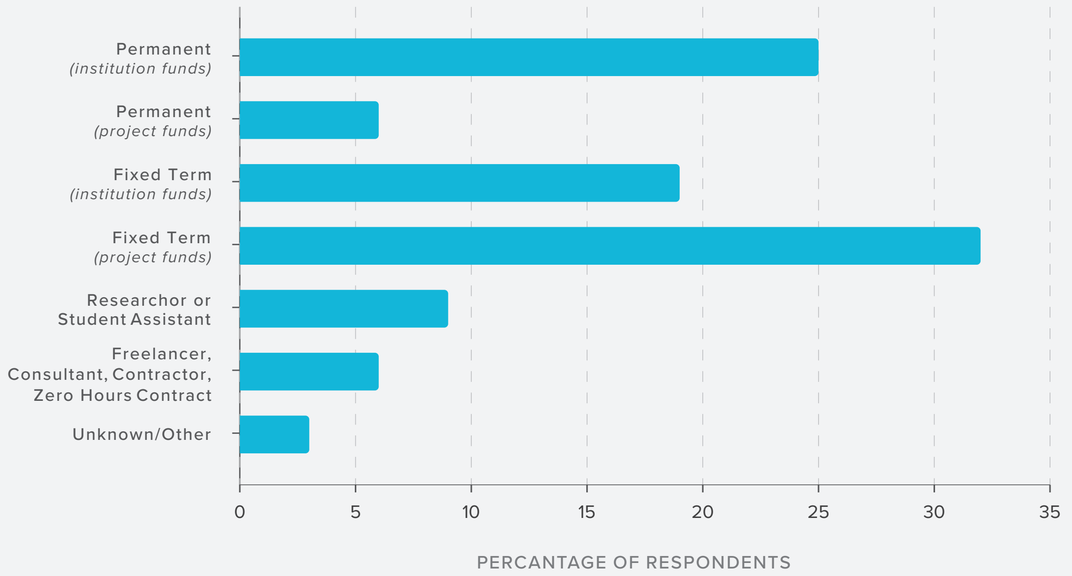

Figure 13: *Staff employment types*

Much has been written on the precarity of careers for those with this capability, and these results show that only 33% of these staff have permanent employment. The National Research Infrastructure Roadmap Exposure Draft notes that the attraction and retention of skilled staff can be challenging, and cites scientific software development as an example of system-wide skills shortages (DESE, 2021).

The same question was asked in the 2018 international RSE survey, which differs from this survey in that it is aimed at individuals who identify as RSEs, whereas this survey asks for responses on behalf of an area. Analysis of the 99 responses from Australian participants (May et al., 2019) are compared to this survey's results in Table 5.

Despite the differences in the focus of the survey respondents for the two surveys, the results have a high degree of similarity, with permanent (institution funds) and fixed term (project funds) being the most common answers in both surveys, and the two lowest ranking responses being research or student assistant, and freelancer, etc. The analysis of the Australian participants in the 2018 RSE survey also contains information on a range of other areas, in relation to employment this includes type of current employer organisation (university, government, national laboratory, etc.) and discipline of work (May et al., 2019).

	RSE SURVEY 2018 (%)	ARDC SURVEY 2021 (%)
PERMANENT (INSTITUTION FUNDS)	37	25
PERMANENT (PROJECT FUNDS)	12	6
FIXED TERM (INSTITUTION FUNDS)	18	19
FIXED TERM (PROJECTS FUNDS)	24	32
RESEARCH OR STUDENT ASSISTANT	5	9
FREELANCER, CONSULTANT, CONTRACTOR, ZERO HOURS CONTRACT	4	7
UNKNOWN/OTHER	n/a	3

Table 5: *Comparison of survey results on employment types*

4.4.3 Knowledge of research software assets

A final question sought to understand what knowledge respondents had of the research software generated by their staff, and asked respondents to list research software developed and/or maintained by their staff that is shared with others, and to ideally provide references to their location (DOIs, GitHub/gitlab/bitbucket repositories, URLs, etc.). 48 of the 70 respondents answered this question, providing a listing of 50 GitHub repositories, 10 bitbucket repositories, R packages, a number of websites, and some comments on bigger efforts, including:

- ~500 scientific software packages we install and upgrade, and tune for the infrastructure.
- We maintain over 300 repositories of software, tools, deployment aids and other related software. Many are open source
- Dozens of bespoke software applications for data capture, reduction, visualisation and analysis, for internal usage by staff and students.

Three respondents also emphasised that they contribute and support other research software that they are not leads on, including making contributions upstream to packages like python libraries, and maintaining branches of code for other organisations.

While it is valuable to understand that at least some research organisations have some awareness of their research software outputs (or those that they contribute to), this is not evidence that the Australian research sector audits research software assets in the same way it now does for research data assets. Open Source Program Offices (OSPOs) are beginning to be integrated in a few US and European universities as a way to improve the management of open source scientific software (Choudhury, 2021). The use of Software Management Plans can also assist organisations (and funders) to understand the volume of research software development occurring, whilst also improving software engineering practices. International disciplinary communities such as ELIXIR in the life sciences encourage use of software management plans (Alves et al., 2021) and research software funders in the Netherlands have set up a working group to explore the possibility of creating a national template for Software Management Plans for Dutch research organisations (Boer, 2021).

5. CONCLUSION

The results of this survey are very useful in understanding the extent to which this research software engineering capability may be meeting Australia's research needs. The analysis indicates there is a sizable and growing community of people working in research software support. It has identified over 6,000 people working in this area nationally, along with over 1,500 people providing training and advice across all levels of skills, and these are probably very conservative estimates. However, the responses to questions such as numbers of staff, sufficiency of staffing, and intention to recruit, along with the large proportion of unfunded staff working in this area strongly indicate the community is not as well-served as it could or should be. The community of people working in software development and support itself needs to be recognised, developed, retained, sufficiently skilled and valued, along with personal development and career path opportunities. The ARDC research software agenda identifies a range of stakeholders who have a role to play in improving recognition of research software and the personnel that support this, and it is important that these stakeholders do so.

Further investigation would be beneficial to provide increased understanding of some areas. Having the survey undertaken regularly would provide clearer longitudinal data and highlight real trends, rather than perhaps showing increased response rates. Gaining more depth of responses from organisations would enhance the coverage and accuracy of the figures. Further clarifying the target respondents may capture both the professional staff and also academic staff working in this area. To achieve this may require greater encouragement of self-assessment, to show its value, but it must be made easy for respondents. For example, the Stanford survey could be utilised by Australian research organisations to gain more insight into their own capability, and may overcome some of the current issues in under-reporting. Administration of this survey by research centres would also help fill the gaps in this survey about this type of stakeholder. Understanding of diversity, equity and inclusivity issues could also improve both research and social outcomes. There are also some areas where international initiatives can provide examples of evolving best practice, such as development of career paths, research software policy, and valuing of research software assets.

If the survey were part of an agreed national framework, with the support of funding bodies and policymakers, it would greatly raise the profile of both the survey and also the issues around the crucial workforce support for all research endeavours. It would also enhance the quality of the measurements, and the ability to interpret the results. The proposed Expert National Research Infrastructure Advisory Group could be relevant given its focus on how to support career pathways, address technical skills shortages and identify capability gaps.

Technology is always advancing, with advances such as quantum and exascale computing set to enable a higher level of performance in research computing that has the potential to rapidly advance research impact. It is crucial that the human capital element of research software infrastructure also achieves the step-changes needed to ensure there are personnel to support this technology.

REFERENCES

ABS. (2020a). Research and Experimental Development, Higher Education Organisations.
https://www.abs.gov.au/statistics/industry/technology-and-innovation/research-and-experimental-development-higher-education-organisations-australia/latest-release

ABS. (2020b, June 19). Research and Experimental Development, Government and Private Non-Profit Organisations, Australia.
https://www.abs.gov.au/statistics/industry/technology-and-innovation/research-and-experimental-development-government-and-private-non-profit-organisations-australia/latest-release

Akhmerov, A., Bazuine, M., Beardsell, J., van den Bogerd, R., Branchett, S., Dunning, A., Keijzer-de Ruijter, M., Cruz, M., Martinez-Lavanchy, P., Schenk, M., Spaargaren, M., & Teperek, M. (2021). TU Delft Research Software Policy.
https://doi.org/10.5281/ZENODO.4629662

Alves, R., Bampalikis, D., Castro, L. J., Fernández, J. M., Harrow, J., Kuzak, M., Martin, E., Psomopoulos, F. E., & Via, A. (2021). ELIXIR Software Management Plan for Life Sciences [Preprint]. BioHackrXiv.
https://doi.org/10.37044/osf.io/k8znb

ARDC. (2019). The Current and Future States of the eResearch Workforce.
https://ardc.edu.au/wp-content/uploads/2019/07/ARDC_InformationSheets.pdf

Barker, M. (2021). ReSA People Roadmap Report.
Zenodo. https://doi.org/10.5281/ZENODO.5633318

Barker, M., Chue Hong, N., Katz, D. S., Leggott, M., Treloar, A., van Eijnatten, J., & Aragon, S. (2021). Research software is essential for research data, so how should governments respond?
https://doi.org/10.5281/zenodo.5762703

Barker, M., Katz, D. S., & Gonzalez-Beltran, A. (2020). Evidence for the importance of research software. Zenodo.
https://doi.org/10.5281/ZENODO.3884311

Barone, L., Williams, J., & Micklos, D. (2017). Unmet needs for analyzing biological big data: A survey of 704 NSF principal investigators. PLOS Computational Biology, 13(10), e1005755.
https://doi.org/10/gcgdcc

Bello, M., & Galindo-Rueda, F. (2020). Charting the digital transformation of science: Findings from the 2018 OECD International Survey of Scientific Authors (ISSA2) (OECD Science, Technology and Industry Working Papers 2020/03).
http://www.oecd.org/digital/charting-the-digital-transformation-of-science-1b06c47c-en.htm

Berente, N., Howison, J., Cutcher-Gershenfeld, J., King, J. L., Barley, S. R., & Towns, J. (2017). Professionalization in Cyberinfrastructure [(SSRN Scholarly Paper ID 3138592). Social Science Research Network].
https://papers.ssrn.com/abstract=3138592

Boer, L. de. (2021, July 9). Transparency in research… And football. Medium.
https://blog.esciencecenter.nl/transparency-in-research-and-football-765a726ab5c9

Buchhorn, M. (2019). Surveying the scale of the research-IT support workforce.
https://ardc.edu.au/wp-content/uploads/2019/07/ARDC-National-Workforce-report-final-v3.pdf

Choudhury, S. (2021). University Open Source Program Offices.
https://ospoplusplus.com/blog/university-open-source-program-offices/

DESE. (2020, December 14). 2021 Research Block Grant Allocations Released [Text]. Department of Education, Skills and Employment; scheme = AGLSTERMS.AglsAgent; corporateName = Department of Education, Skills and Employment; address = 50 Marcus Clarke St, Canberra City, ACT 2601; contact = +61 1300 566 046.
https://www.dese.gov.au/research-block-grants/announcements/2021-research-block-grant-allocations-released

DESE. (2021). National Research Infrastructure Roadmap Exposure Draft.
https://2021nriroadmap.dese.gov.au/wp-content/uploads/2021/11/Draft-2021-NRI-Roadmap-29-NovPM-FINAL.pdf

REFERENCES

DORA. (2021). Signers.
https://sfdora.org/signers/

Duca, D. (2019). Using and developing software in social science and humanities research.
https://ocean.sagepub.com/blog/using-and-developing-software-in-social-science-and-humanities-research

Eghbal, N. (2016). Roads and Bridges: The Unseen Labor Behind Our Digital Infrastructure.
https://www.fordfoundation.org/media/2976/roads-and-bridges-the-unseen-labor-behind-our-digital-infrastructure.pdf

EOSC. (2021). Strategic Research and Innovation Agenda of the European Open Science Cloud (Version 1.0 15).
https://www.eosc.eu/sites/default/files/EOSC-SRIA-V1.0_15Feb2021.pdf

EOSC Executive Board. (2020). Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC) (Version 0.8). EOSCSecretariat.eu.
https://www.eoscsecretariat.eu/sites/default/files/eosc-sria-v08.pdf

European Commission. (2017). Evaluation of Research Careers fully acknowledging Open Science Practices; Rewards, incentives and/or recognition for researchers practicing Open Science,. Publications Office of the European Union.
http://dx.doi.org/doi: 10.2777/75255.

European Commission. (2019). Open Science.
https://ec.europa.eu/info/research-and-innovation/strategy/goals-research-and-innovation-policy/open-science_en

G7. (2021). G7 Research Compact.
https://www.g7uk.org/wp-content/uploads/2021/06/G7-2021-Research-Compact-PDF-356KB-2-pages-1.pdf

Hettrick, S. (2014). It's impossible to conduct research without software, say 7 out of 10 UK researchers.
https://www.software.ac.uk/blog/2014-12-04-its-impossible-conduct-research-without-software-say-7-out-10-uk-researchers

Hettrick, S., Antonioletti, M., Carr, L., Chue Hong, N., Crouch, S., De Roure, D., Emsley, I., Goble, C., Hay, A., Inupakutika, D., Jackson, M., Nenadic, A., Parkinson, T., Parsons, M. I., Pawlik, A., Peru, G., Proeme, A., Robinson, J., & Sufi, S. (2014). UK Research Software Survey 2014 [Data set]. Zenodo.
https://doi.org/10.5281/zenodo.14809

Honeyman, T., & Treloar, A. (2021). A National Agenda for Research Software.
https://doi.org/10.5281/ZENODO.4940273

Japan Science and Technology. (2017). Strategy.
https://www.jst.go.jp/EN/about/strategy.html

Katz, D. S., McHenry, K., Reinking, C., & Haines, R. (2019). Research Software Development & Management in Universities: Case Studies from Manchester's RSDS Group, Illinois' NCSA, and Notre Dame's CRC. 2019 IEEE/ACM 14th International Workshop on Software Engineering for Science (SE4Science), 17–24.
https://doi.org/10/ggx4zz

Knowles, R., Mateen, B. A., & Yehudi, Y. (2021). We need to talk about the lack of investment in digital research infrastructure. Nature Computational Science, 1(3), 169–171.
https://doi.org/10.1038/s43588-021-00048-5

May, N., Lohani, N. P., & Sinha, M. (2019). Research Software Engineers: Australian Results of the 2018 Survey.
https://rse-aunz.github.io/assets/RSE_Survey_2018_AU_v3.pdf

Mons, B. (2020). Invest 5% of research funds in ensuring data are reusable. Nature, 578(7796), 491–491.
https://doi.org/10.1038/d41586-020-00505-7

Nangia, U., & Katz, D. S. (2017). Track 1 Paper: Surveying the U.S. National Postdoctoral Association Regarding Software Use and Training in Research.
https://doi.org/10.6084/m9.figshare.5328442.v1

Nisbet, S., & Gray, M. (2018). Supporting bioinformatics: (p. 43).

REFERENCES

OECD. (2019). Measuring the Digital Transformation: A Roadmap for the Future. OECD.
https://doi.org/10.1787/9789264311992-en

OECD. (2020). Building digital workforce capacity and skills for data-intensive science (OECD Science, Technology and Industry Policy Papers No. 90; OECD Science, Technology and Industry Policy Papers, Vol. 90).
https://doi.org/10.1787/e08aa3bb-en

Office of the Chief Science Advisor of Canada, I. (2020, February). Roadmap for Open Science.
https://www.ic.gc.ca/eic/site/063.nsf/eng/h_97992.html

Scroggins, M. J., & Pasquetto, I. V. (2020). Labor Out of Place: On the Varieties and Valences of (In)visible Labor in Data-Intensive Science. Engaging Science, Technology, and Society, 6, 111–132.
https://doi.org/10.17351/ests2020.341

SPARC Europe, & Digital Curation Centre. (2019). An Analysis of Open Science Policies in Europe v4.
https://doi.org/10.5281/ZENODO.3379705

Springer, R. (2019). Counting Data Librarians.
https://sr.ithaka.org/blog/counting-datalibrarians/ (accessed

Stanford Software Survey. (2020).
https://stanford-rc.github.io/stanford-software-survey

The Carpentries. (2021). Past_workshops.json.
https://feeds.carpentries.org/

APPENDIX A: SURVEY QUESTIONS

Research software capability in Australia

Survey information

The Australian Research Data Commons (*ARDC*) is conducting a survey to help understand the scale and distribution of software engineering or development capability within research organisations in Australia. This is a capability that is currently spread across professional and academic roles.

We aim to obtain an accurate indication of the size of the national workforce involved in supporting research software. The survey supports the implementation of the ARDC's *National Agenda for Research Software in Australia.*

Who should complete the survey?

We are seeking responses from across the Australian research sector, including research institutions and supporting eInfrastructure providers, such as NCRIS facilities, universities, research centres and institutes, Publicly Funded Research Agencies (PFRAs), and regional/state/national eInfrastructure organisations.

Group/Project Leads, (Senior) Managers, Chief/Principal Investigators or Academic Leadership should complete the survey on behalf of their unit, but in some cases it may be appropriate for individuals to respond on behalf of their unit or team.

The survey can be completed multiple times, or combine responses over a number of areas. If we receive separate responses from one organisation, we will appropriately combine the responses.

What does the survey ask?

The survey asks you to provide information about research software professionals at your institution, organisation, or group as appropriate.

The survey contains 17 questions about the size and capability of your group and 3 questions about research software management and career progression pathways at your institution.

The questions are designed to be easy for quick, estimated responses. The survey should take between 5-10 minutes to complete. There are also spaces to provide additional information or clarification.

Please share the survey with other key organisational leaders that you are aware of with research software development activities.

If you have any questions about the survey please *contact us.* At the end of the survey, we will invite you to register to receive the outcomes of this survey.

Privacy and confidentiality

All survey data will be de-identified for inclusion in the final report. The data will be collected and handled in accordance with the *ARDC privacy policy.*

Please complete this survey by COB Friday 19 November 2021.

APPENDIX A: SURVEY QUESTIONS

Section 1: Respondent and organisational details

This section helps us (i) classify your response to compare with others, and (ii) gives us a contact for us to come back if we need any clarification.

PLEASE NOTE: All survey data will be de-identified for inclusion in the final report.

1. Your name? (This detail helps us to identify and resolve duplicate submissions only, and is not taken as consent to contact you)
2. [1]Organisation/group/centre, etc. that you are responding for? (This is used to understand the type of organisation/group, and to ensure that duplicate or overlapping responses are handled appropriately.) If multiple then please submit a separate response for each, to help with the analysis
3. What are the two-digit Fields of Research (FoR) code/s of the responding area (choose all that are relevant)?

- ☐ 30 AGRICULTURAL, VETERINARY AND FOOD SCIENCES
- ☐ 31 BIOLOGICAL SCIENCES
- ☐ 32 BIOMEDICAL AND CLINICAL SCIENCES
- ☐ 33 BUILT ENVIRONMENT AND DESIGN
- ☐ 34 CHEMICAL SCIENCES
- ☐ 35 COMMERCE, MANAGEMENT, TOURISM AND SERVICES
- ☐ 36 CREATIVE ARTS AND WRITING
- ☐ 37 EARTH SCIENCES
- ☐ 38 ECONOMICS
- ☐ 39 EDUCATION
- ☐ 40 ENGINEERING
- ☐ 41 ENVIRONMENTAL SCIENCES
- ☐ 42 HEALTH SCIENCES
- ☐ 43 HISTORY, HERITAGE AND ARCHAEOLOGY
- ☐ 44 HUMAN SOCIETY
- ☐ 45 INDIGENOUS STUDIES
- ☐ 46 INFORMATION AND COMPUTING SCIENCES
- ☐ 47 LANGUAGE, COMMUNICATION AND CULTURE
- ☐ 48 LAW AND LEGAL STUDIES
- ☐ 49 MATHEMATICAL SCIENCES
- ☐ 50 PHILOSOPHY AND RELIGIOUS STUDIES
- ☐ 51 PHYSICAL SCIENCES
- ☐ 52 PSYCHOLOGY
- ☐ Other

4. Number of researchers in this area (including postgraduate students)?

- ☐ 0
- ☐ 1-5
- ☐ 5-10
- ☐ 10-20
- ☐ 20-50
- ☐ 50-100
- ☐ 100-200
- ☐ 200-500
- ☐ 500-1,000
- ☐ 1,000-5,000
- ☐ 5,000-10,000
- ☐ 10,000+

---

1 ■ Denotes a question where an answer was compulsory

APPENDIX A: SURVEY QUESTIONS

Section 2: Research software staff and team composition

The following questions are about staff with computing skills who support researchers in the development/engineering of software tools to facilitate the generation of research outcomes; and researchers/students who undertake such work noting it may not be a formal part of their job description. The roles of these staff could include coding, documentation, project management, software maintenance, but does not include research software trainers, business analysts, helpdesk staff, etc.

We want to count people who primarily work with researchers, and may include employees, contractors, interns, researchers and students. The most important question is the 'people count'. Please try to also provide equivalent-full-time (EFT) totals to allow for better comparisons across the sector. The 'unfunded count' aims to identify people who are doing this work but it is not their main responsibility (like researchers and students).

Please provide your best estimates for each of the questions below, and err on the side of including, rather than excluding people in your counts. The questions are designed to be easy for quick estimated responses. There is space to provide additional information or clarification.

5. How many (people/EFT/unfunded) develop/engineer/maintain software used in research analysis?

	0	1-5	5-10	10-20	20-50	50-100	100-200	200-500	500+
PEOPLE COUNT									
EFT									
UNFUNDED COUNT									

6. How sufficient is this number to meet organisational needs?
   - ☐ Sufficient
   - ☐ Nearly sufficient
   - ☐ Slightly inadequate
   - ☐ Severely inadequate
7. Do you plan to recruit more such people in the short-medium term? (1-3 years)
   - ☐ Yes
   - ☐ Maybe
   - ☐ No
8. Approximately what percentage are employed through each of the following arrangements? (numbers must sum to 100)
   - ☐ Permanent (institution funds)
   - ☐ Permanent (project funds)
   - ☐ Fixed term (institution funds)
   - ☐ Fixed term (project funds)
   - ☐ Research or student assistant
   - ☐ Freelancer, consultant, contractor, zero hours contract
   - ☐ Unknown/other
9. What are the most common job titles for these staff?

10. How many different levels of research software staff with different responsibilities are there? (e.g. are there junior and senior roles, team leaders, supervisors, managers, etc.?)

☐ 1 ☐ 3 ☐ 5+

☐ 2 ☐ 4 ☐ NA

11. Please feel free to add any additional information or clarification on the preceding questions, or any considerations we did not ask about.

12. Do your research software staff have the skills they need?

☐ Yes ☐ Maybe ☐ No

13. Do your research software staff have mechanisms to acquire the skills they need?

☐ Yes ☐ Maybe ☐ No

14. How do/could your research software staff access training in this area? This is deliberately an open-ended question, to collect a variety of approaches.

15. How many people provide advice, education/training to researchers/research teams on relevant tools and applications?

	0	1-5	5-10	10-20	20-50	50-100	100-200	200-500	500+
PEOPLE COUNT									
EFT									
UNFUNDED COUNT									

16. Please feel free to add any additional information or clarification on the preceding questions (Questions 13-16).

Section 3: Research software

This section aims to raise awareness of locally developed, maintained or supported research software. Research software can be defined as including domain-specific tools, and scientific infrastructure, and excludes general-purpose software, non-scientific infrastructure, and operating systems.

17. Please list research software developed and/or maintained by your staff that is shared with others. Ideally provide references to their location. (DOIs, GitHub/gitlab/bitbucket repositories, URLs, etc., are all acceptable.)

18. What, if any, policy does your organisation have that relates to research software? (E.g. open science policy, software management plan policy, etc.)

19. Does your organisation recognise software development or maintenance effort towards academic progression? If so, how?

20. A contact email address? (Only to be used if further clarification is needed, or to fulfil choices in the next question)

21. Would you like to know more?

☐ Please notify me of the outcomes of the survey

☐ Please subscribe me to the ARDC newsletter

☐ Please contact me to discuss the "National Agenda for Research Software"

Thank you very much for your participation!

APPENDIX B: JOB TITLES

80 unique job titles were identified by survey respondents.

Job Title	Number of Mentions
Academic	1
Analyst Programmer	1
Assistant Lecturer	1
Associate Lecturer	2
Associate Professor	2
Computational Scientist	1
Data Architect	1
Data Engineer	1
Data Management Specialist	1
Data Manager	2
Data Science Software Engineer	1
Data Scientist	1
Data Warehouse Developer	2
Developer	3
DevOps Engineer Junior	1
DevOps Engineer Senior	1
DevOps Systems Developer	1
Director of Operations	1
Doctor	1
eResearch Analyst	1
eResearch Team Staff	1
Experimental Scientist	1
High Performance Computing Specialist	1
Informatics Fellow	2
Informatics Software Developer	1
Information Services Manager	1
IT Assistant	1
Junior Developer	1
Junior Software Developer	1
Lecturer	5
Library Research Support	1
NLP Engineer	1
PhD Student	7
Postdoctoral Research Associate	1
Postdoctoral Researcher	8
Practice Lead	1
Professor	4
Project Manager	1
Project Officer	1
Research Analytics and Data Specialist	1
Research Assistant	8
Research Associate	2
Research Data and Analytics Advisor	1
Research Data Consultant	1
Research Data Steward	1
Research Data Systems Manager	2
Research DevOps	1

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
ARDC is enabled by NCRIS

APPENDIX B: JOB TITLES

Job Title	Number of Mentions
Research Engineer	1
Research Engineer Group Lead	1
Research Fellow	1
Research Manager	1
Research Officer	1
Research Projects Officer	1
Research Scientist	3
Research Software Developer	1
Research Support	2
Researcher	1
Research Software Engineer	2
Scientific Programmer	2
Scientist	1
Senior Bioinformatician	1
Senior Data Scientist	1
Senior Development Engineer	1
Senior Lecturer	2
Senior Research Fellow	1
Senior Research Informatics Technical Officer	1
Senior Research Officer	1
Senior Software Developer	1
Senior Software Engineer	4
Software Developer	6
Software Engineer	4
Software Engineer (Research)	1
Solutions Architect	1
Statistical Consultant	1
System Architect	1
Systems Administrator	2
Systems Administrator CRM	1
Technical Officer	1
Web Developer	2
Wipro Consultant DITM	1

ACKNOWLEDGEMENTS

We would like to acknowledge and thank the members of the *Institutional Underpinnings* Policy Working Group for their input to this document.

FEEDBACK

We welcome your feedback on this guide. Please email *contact@ardc.edu.au* with any comments or questions.

ABOUT THE AUSTRALIAN RESEARCH DATA COMMONS

The Australian Research Data Commons (ARDC) enables the Australian research community and industry access to nationally significant, data intensive digital research infrastructure, platforms, skills and collections of high quality data.

The ARDC is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS).

ARDC is enabled by NCRIS

Australian Research Data Commons

CONTACT

ardc.edu.au
+61 3 9902 0585
contact@ardc.edu.au

FOLLOW

@ardc_au
australian-research-data-commons
subscribe to our newsletter

NCRIS
National Research Infrastructure for Australia
An Australian Government Initiative

The ARDC is enabled by NCRIS