



International  
Centre for  
Radio  
Astronomy  
Research



**PAWSEY**  
supercomputing centre

# AusSRC Radio Astronomy Data User Community Survey

Slava Kitaeff, Louisa Quartermaine

# Survey Rationale

- Part of community consultations for AusSRC development
- Build on the information provided by the previously collected data during AusSRC workshops (2017, 2019)
- Focus on identifying current gaps in the services, support and infrastructure
- How this information could be mapped onto the community expectations from the AusSRC?
- Update the definition and requirements of AusSRC
- Next version of AusSRC white paper and Business Plan.

# Survey Structure

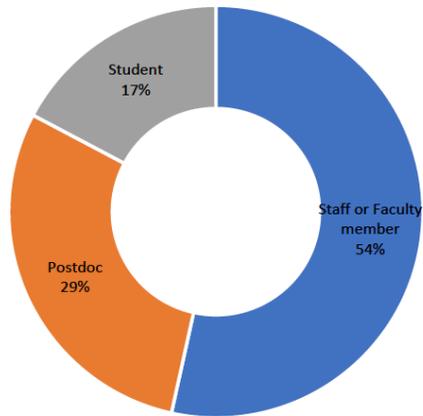
1. **User Profiling** - this survey section sought information about a user of radio astronomy data, such as areas of scientific interest, position and role(s) and skill level working with radio data as well as various computing skill levels.
2. **Projects Profiling** - this survey section sought information on the projects with which users are involved, such as type of projects, nature of the collaboration, radio astronomy facility used. In this section we have also tried to identify some of the common observational project related challenges.
3. **Data Profiling** - this survey section sought information about the data and some specific challenges related to it.
4. **Software and Tools Profiling** - this survey section sought information on the type of software and tools used to process and interact with radio astronomy data, such as data reduction, statistical analysis, cross matching of catalogues, visualisation and image processing, as well as identifying the issues encountered when using those software and tools.
5. **Data Processing Profiling** - this survey section sought information on where users process their data, what type of data manipulation and interaction they perform, their satisfaction with data and compute facilities used, as well as the issues they encountered when processing data.
6. **Expectations from AusSRC** - this section requested users to rate the importance of what AusSRC should be providing in the future.

# Conducting the Survey

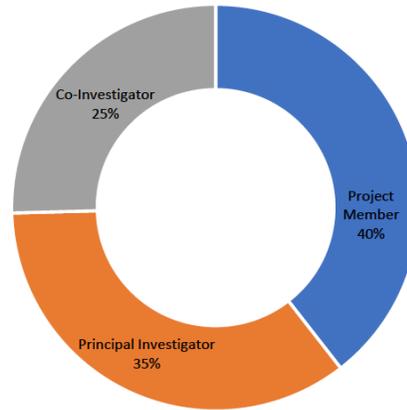
- October-November 2019
- Qualtrics survey software (provided by UWA)
- 33 questions
- Distributed through Astronomical Society of Australia (twice, 6 weeks apart)
- **Seventy-four** survey responses were received (~60 to 70% of the member of Australian radio astronomy science community?).

# Users and survey confidence

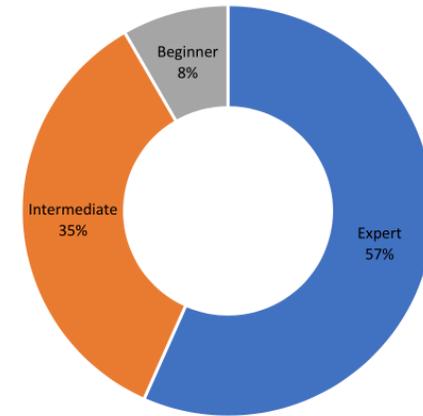
6.1.2. What position do you hold?



6.1.3. What project role do you hold?

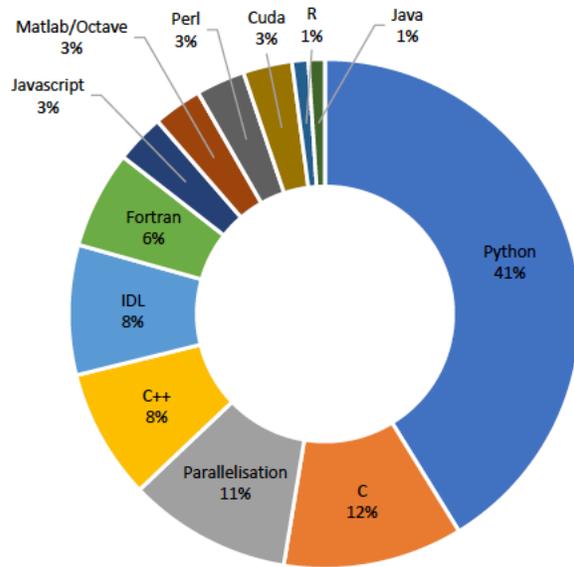


6.1.4. What is your skill level working with radio astronomy data?

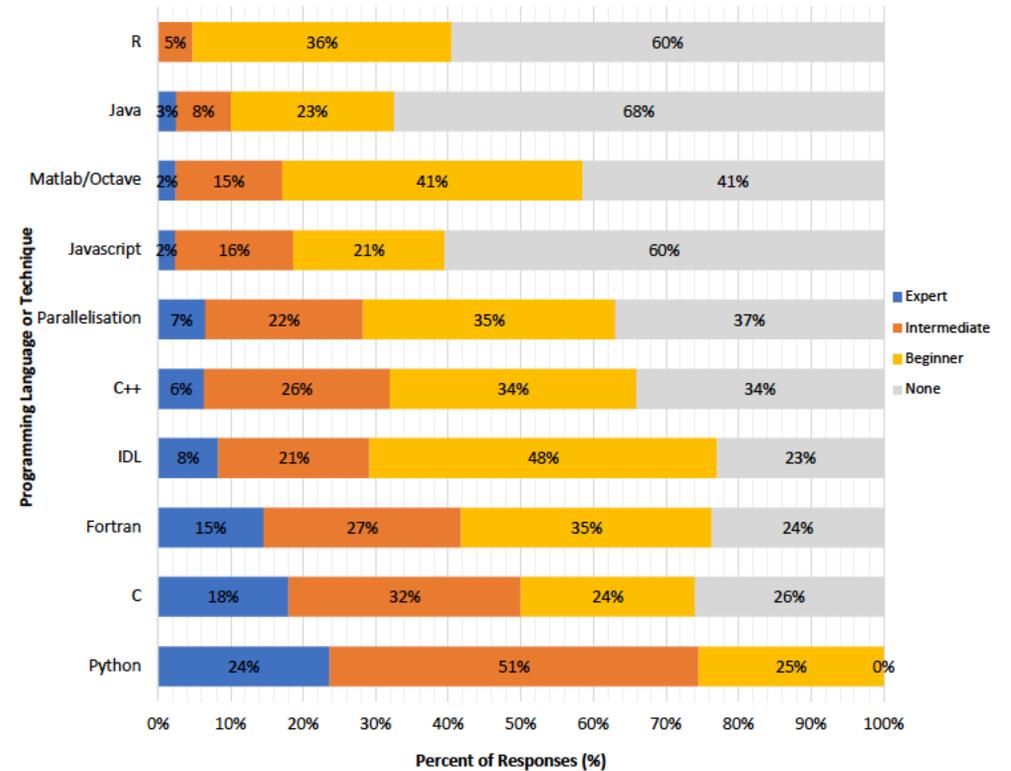


# Programming Skills

6.1.5. Which programming languages and techniques do you use often?

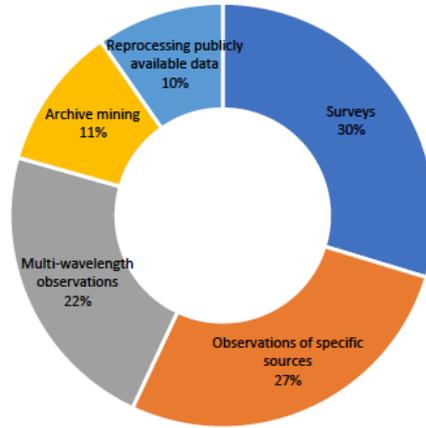


6.1.6. What is your skill level with these languages and techniques?

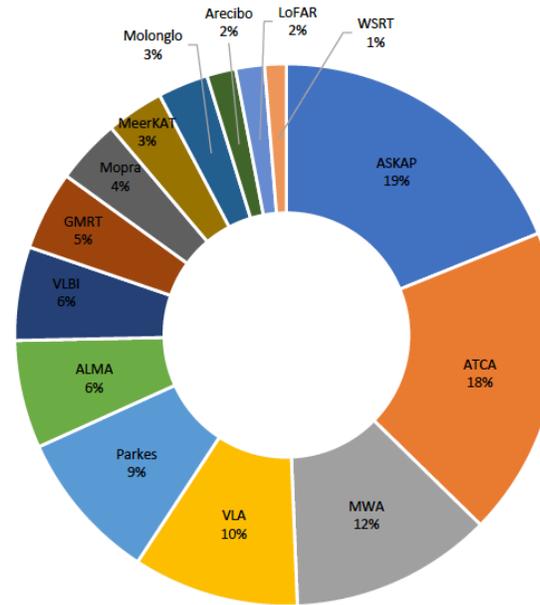


# Projects

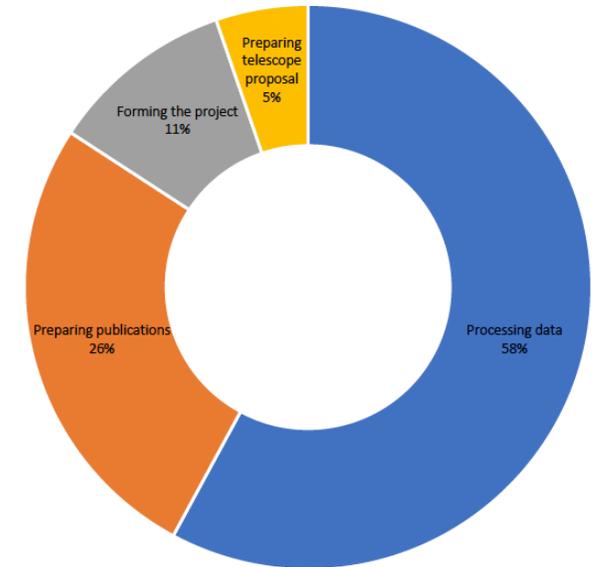
6.2.1. With what types of projects are you currently involved?



6.2.3. Which radio astronomy facilities do you currently use?

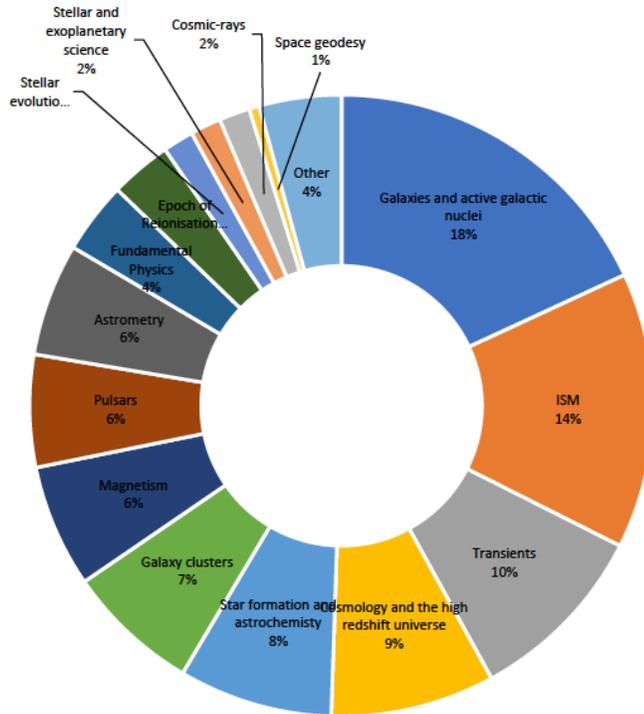


6.2.5. What is the most challenging stage in observational projects?

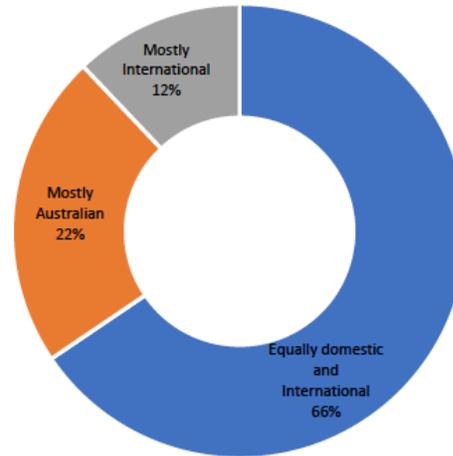


# Users and survey confidence: engagement in science and SKA

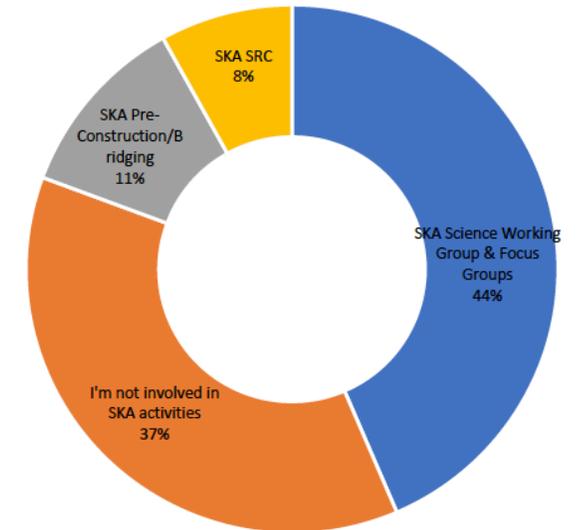
## 6.1.1. What are your areas of scientific interest



## 6.2.2. What is the nature of your current collaborations?

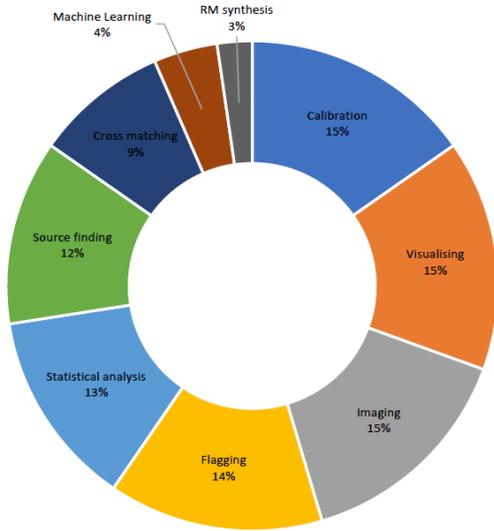


## 6.1.7. Do you participate in any SKA related activities?

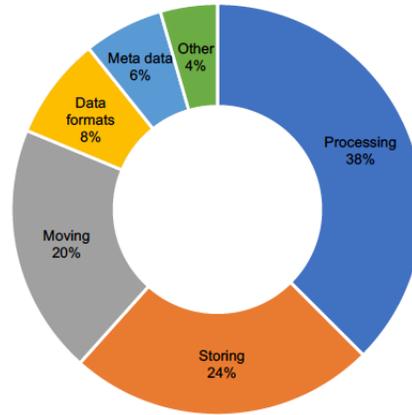


# Data

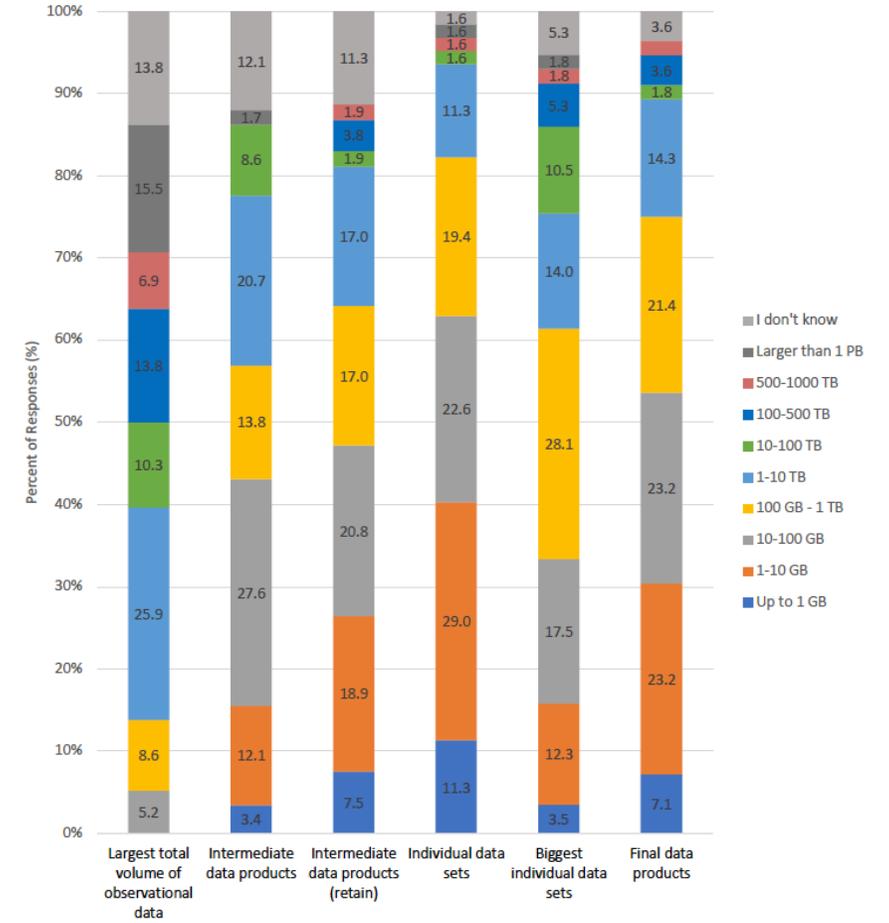
6.5.2. What type of data manipulation/interaction do you perform?



6.3.3. What are the biggest issues you encounter when working with radio data?

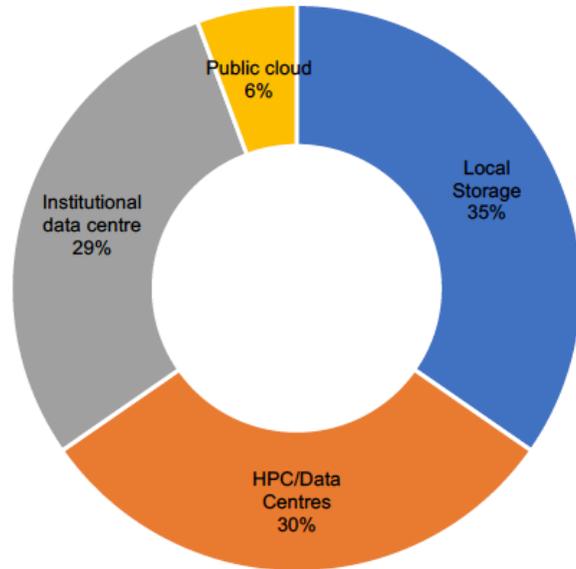


6.3.1. What is the size of the data with which you interact?

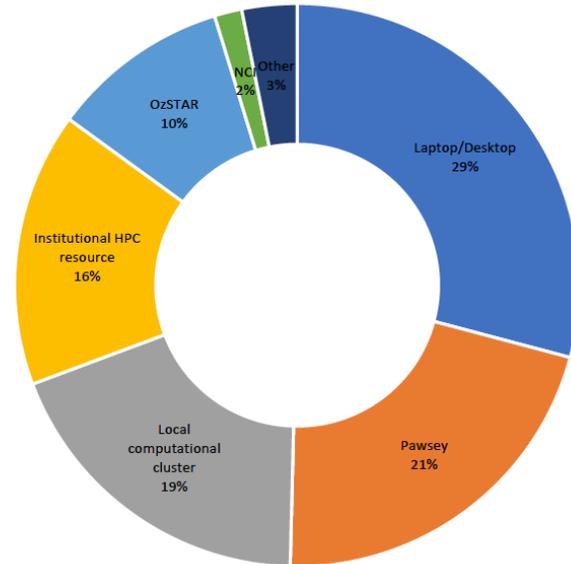


# Data

6.3.2. Where is your data stored?



6.5.1. Where do you process your data?



# Issues working with the data

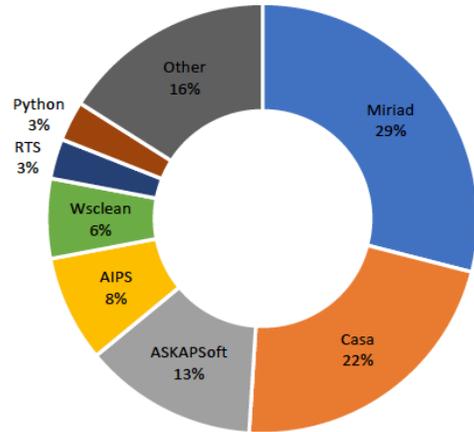
## 6.3.4. Biggest issues working with the data – elaborated

The respondents were asked to elaborate on what they considered as the biggest issues encountered when working with radio astronomy data. The individual responses can be summarised into the following most common issues:

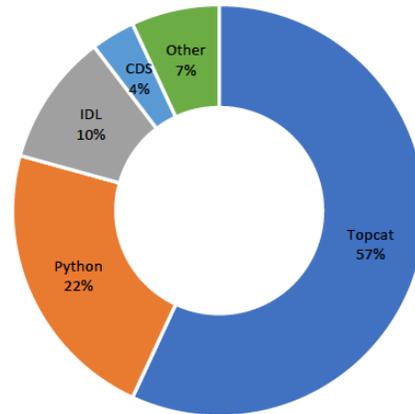
- The large volume of data requires the use of supercomputers which have become bottlenecks in the areas of processing, queue lengths, disk space, retrieving data and are considered a significant impediment to scientific research;
- The large amounts of data require more storage which is often not available, is limited and is not long lived;
- Moving and accessing large amount of data is problematic - time consuming and challenging;
- Manipulating, interacting and visualising large data sets can be problematic;
- The lack of meta data is a problem for querying archives;
- There is a lack of information available on the quality of data which leads to labor intensive and time-consuming manual inspection of data.

# Software and Tools

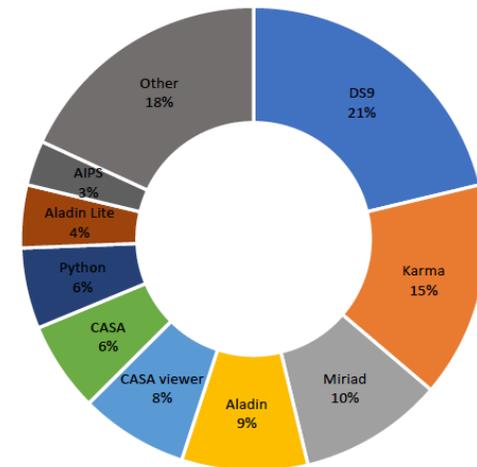
6.4.1. What data reduction software do you use?



6.4.2. What software do you use to work with catalogues for statistical data analysis and cross matching?



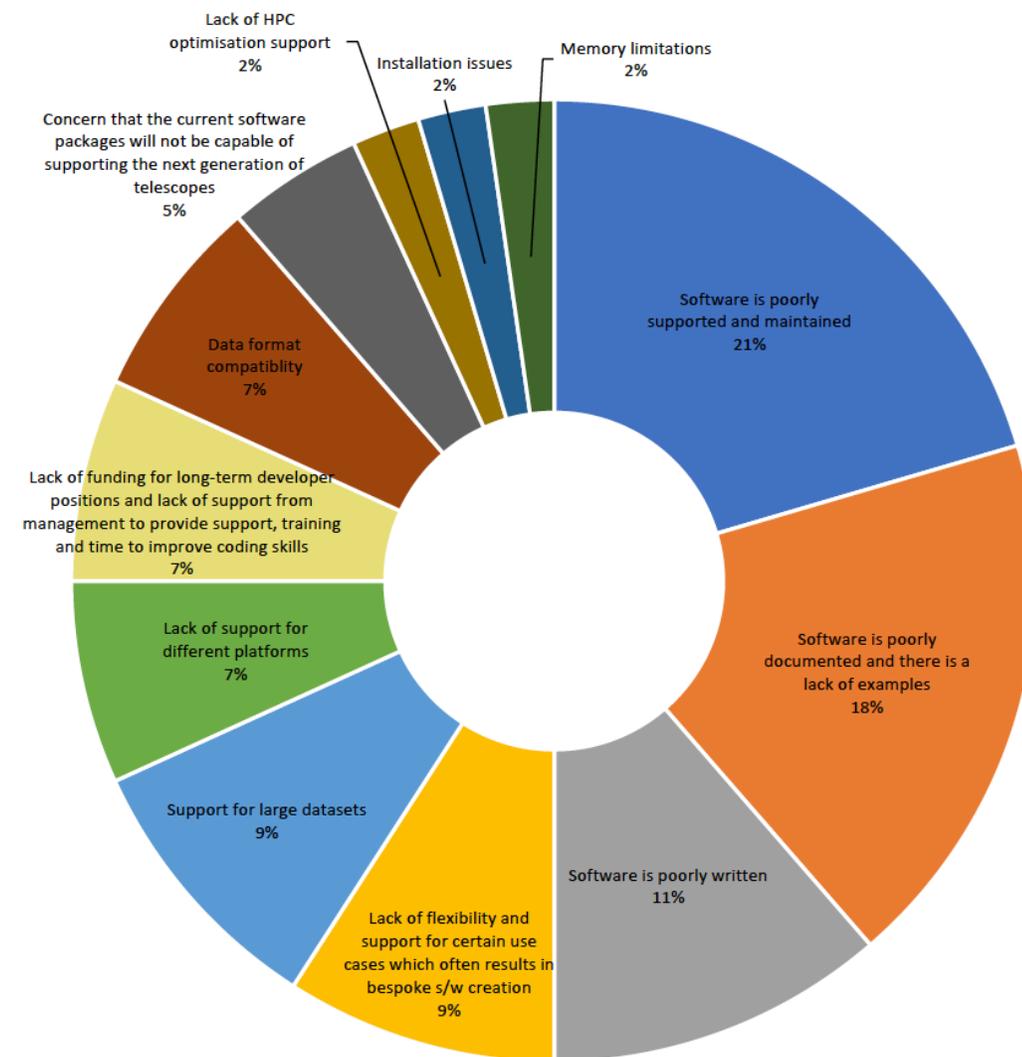
6.4.3. What software do you use for visualisation and image processing?



# Issues with the tools

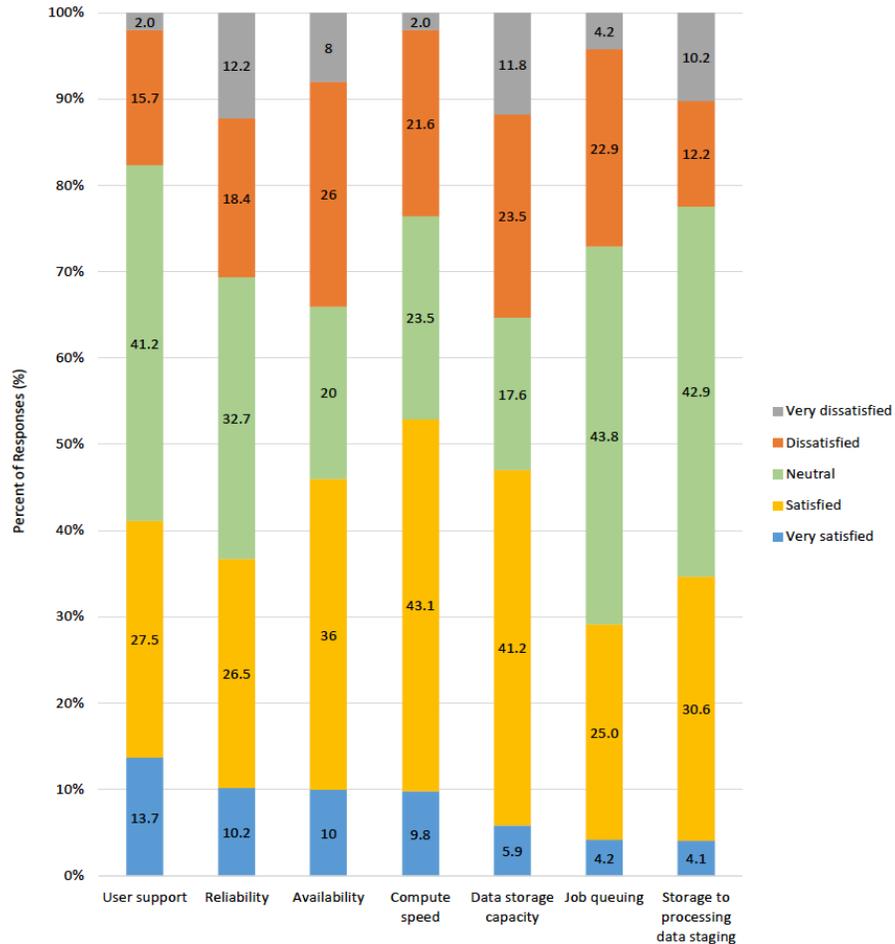
What are the biggest issues that you encounter using astronomy software and tools?

- Software is poorly written;
- Software is poorly documented and there is a lack of examples;
- Software is poorly supported and maintained – often issues are encountered when upgrading to newer versions;
- Concern that the current software packages will not be capable of supporting the next generation of telescopes;
- Lack of support for different platforms;
- Lack of flexibility and support for certain use cases often results in bespoke software creation;
- Lack of HPC optimization support;
- Lack of funding for long-term developer positions and lack of support from management to provide support, training and time to improve coding skills.
- No support for large datasets
- Data format compatibility issues
- Installation issues
- Memory limitations

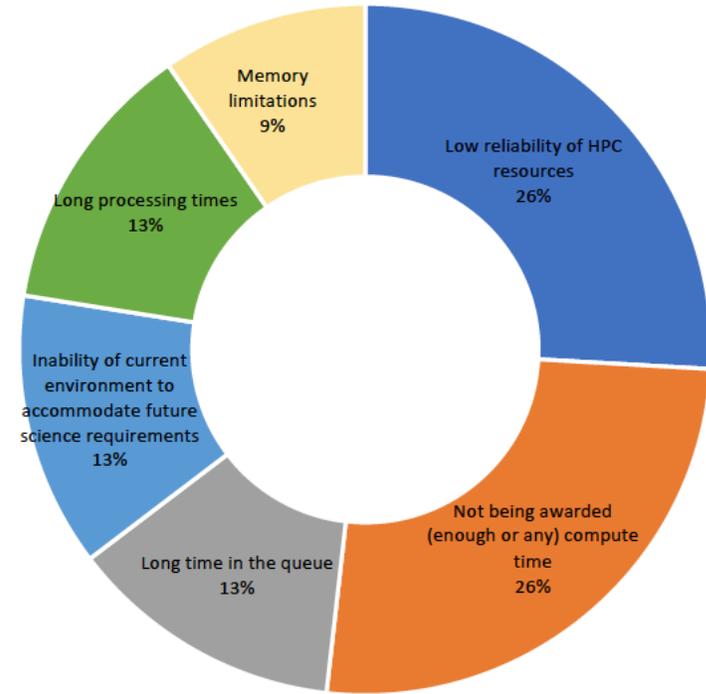


# Issues with facilities

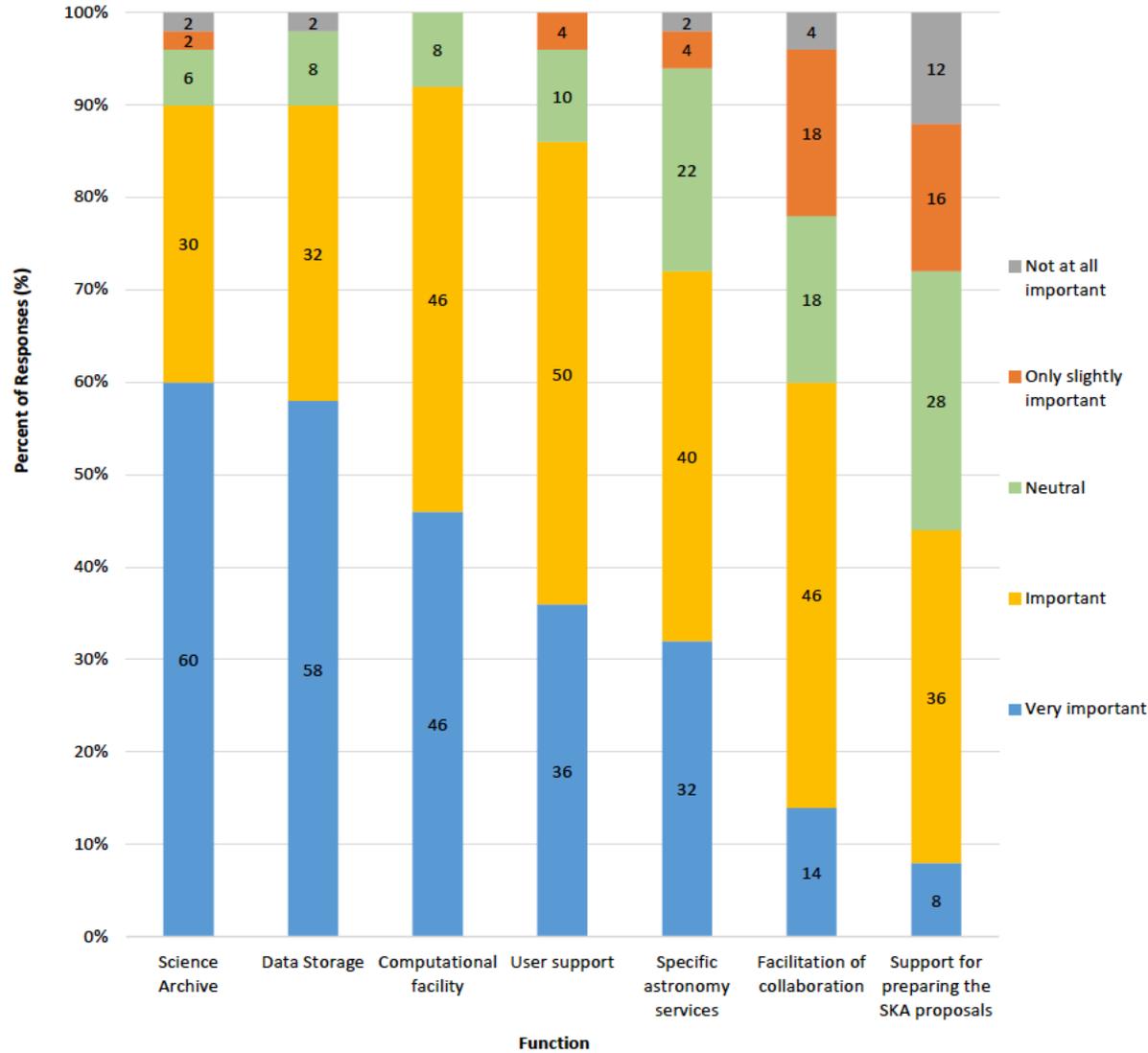
6.5.3. What is your level of satisfaction with the data and compute facilities you use?



5.4. What are the biggest issues you encounter when processing the data?



# Expectations of AusSRC



# Not quite conclusions (observations)

- Survey respondents had good connections to SKA activities, which allows confident (more or less) linkage between the findings and the objectives and requirements for the future Australian SKA Region Centre and its current design study activities.
- The survey indicates the current existence of a broad range of types of observational projects, as well as a broad range of telescope engagement, both nationally and internationally.

# Conclusions: Gaps & Challenges

The survey indicates several gaps and challenges in making science with today's radio data:

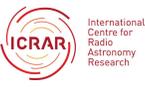
- Processing is a significant challenge across all stages of the data path, including, moving and storing large volumes of data, as well as the computational reduction of this data.
- HPC support of existing software is minimal, and maintenance of the software is often poor.
- More than 2/3 of the users are using HPC facilities to reduce data; however, the lack of dependable/dedicated HPC is a significant issue.

# Conclusions: AusSRC

- The survey reveals the expectation in the community that the SRCs are to be an operational resource for the SKA Archive, Store, and processing needs.

# Conclusions

- This survey is a snapshot of the current state
- With the ASKAP surveys now underway, the development of new software tools, and the refreshed infrastructure at Pawsey, the focus and challenges may shift over time
- AusSRC Design Study Program needs to track and follow such changes; therefore, we anticipate taking another such a snapshot survey in November 2021.



PAWSEY  
supercomputing centre

# Full report

- <https://aussrc.org/wp-content/uploads/2020/07/AusSRC-2019-Radio-Data-User-Survey-Report.pdf>