



Response to the CASA Users Committee Annual Report 2019

Change Record

| Version | Date | Reason |
|---------|------|---------------|
| 0.1 | | Initial Draft |
| | | |

Table of Contents

| | |
|--|----------|
| Change Record | 1 |
| Table of Contents | 2 |
| Overview | 3 |
| 2019 CUC Recommendations and Responses | 3 |
| 6.2 User Servicing & Outreach | 3 |
| User servicing | 3 |
| Outreach | 4 |
| 6.3 CARTA | 6 |
| 6.4 Transition from CASA 5.x to CASA 6.x | 6 |
| 6.5 Specific functionality: polarization and imaging | 7 |
| Polarization | 7 |
| Imaging | 7 |
| 6.6 Archival Data Processing, Pipelines, and SRDP | 8 |
| 6.7 Documentation: CASAdocs and CASA Guides | 10 |
| 6.8 CASA Reliability & Performance | 11 |
| 6.9 Support for Other telescopes | 13 |
| 6.10 CASA Next Generation Infrastructure (CNGI) | 14 |

Overview

This document is the official response from the Common Astronomy Software Applications (CASA) group to the recommendations of the 2019 CASA Users Committee (CUC). The recommendations of the CASA Users Committee along with the full context of Committee discussion are contained in the Annual report dated 31 December 2019.

We wish to thank the CASA Users committee for their time and effort. The recommendations in their annual report are constructive and beneficial to informing CASA strategic goals and decision making. The CASA development team agrees with these recommendations in most instances and will strive to satisfy the commitments in our response over the next year. We are limited in some cases by available resources but generally feel that we can provide an acceptable response or mitigation to every CUC request.

Below we discuss the recommendations that the committee has explicitly called out in their document. We have maintained the section numbering from the CUC annual report to simplify cross-referencing between the documents. In this document we have extracted and summarized the recommendations of the users committee (in black) from their report and provided the response of the CASA project (in blue).

2019 CUC Recommendations and Responses

6.2 User Servicing & Outreach

User servicing

1. Although the number of new Helpdesk tickets matched the number of solved tickets, it was unclear if there is a backlog of tickets which have remained open for longer periods, and if so what the causes are. The committee would like to see a breakdown of open tickets including the number which have been open for a year or more and are not due to unresponsive users (which is obviously beyond the helpdesk's control). Obviously, if there are a significant number of such long-standing tickets, addressing them should be a priority.

We will attempt to extract that information from the NRAO and ALMA Helpdesk systems, to determine how many and which tickets that are under our control have been open for over a year. The User Liaison, in consultation with the CUC, will prioritize any open tickets related to CASA bug reports or feature requests from the user community. Where relevant, the User Liaison will present these tickets for further prioritization with all other

stakeholders at the CASA stakeholders meeting for implementation in the next CASA release.

2. Telemetry reporting is proving to be very useful, but some questions arose as to the validity of the numbers. Between the CUC members we have crashed several of the tasks more frequently than reported. Also note that the message “CrashReporter initialized” is confusing for new CASA users. Users would benefit from more informative messages or documentation of existing ones.

We will investigate crash reporter anomalies and continue to refine the collection process. With the migration to CASA 6, some aspects of telemetry and crash reporting will need to be updated and reevaluated in light of the new usage patterns expected.

3. The spring meeting between CUC chair and CASA liaison was a success and should be continued.

We agree and will arrange another spring telecon with the CUC also this year.

4. User input for the CASA team is now much better structured via a few limited channels. Advertising these channels more clearly on the CASA website will help users find the right channel.

We will update the CASA website to better advertise and explain the channels through which users can contact the CASA team.

Outreach

1. The CUC agrees with the CASA team that the need for an ongoing user survey is limited. With the feedback from the telemetry reports we recommend the survey be discontinued. Once the move to CASA6 is complete, a new survey could be considered.

We will discontinue the CASA users survey, and consider a new survey in the future.

2. We are glad to see that last year’s recommendation to make outreach materials was followed up. This enables essentially anyone to advertise CASA at meetings around the world, thus removing the intrinsic US bias. A small amount of coordination can help to motivate people to actively advertise CASA at the main continental meetings such as AAS and EAS. As it is not clear where the materials can be found, a link on the CASA homepage would be helpful.

We have in the past coordinated CASA advertisements for major conferences in which

CASA team members or NRAO colleagues attended. Thus far this has mostly focused on the AAS and ADASS conferences. The User Liaison considered attending the EAS in 2020, but the EAS organizing committee decided to have a pure science focus this year, hence advertising software does not fit any of the topics.

We would welcome CUC members to advertise CASA at conferences and meetings. The User Liaison is available to coordinate this and update materials that CUC members or others wish to show at meetings or conferences.

The existing outreach materials are linked from the CASA website, but we will update the website to make those links clearer.

3. The memo series is an excellent step in providing in-depth information. This series was new to the CUC and could use better advertising to the users (it is not found on the CASA website). The memo series has the potential to overlap with the Newsletter, though for now they clearly serve a distinct purpose.

The CASA Memo Series is indeed new, and linked from within CASA Docs. The reason to host the Memo Series on CASA Docs is that the memos give in-depth information mostly on things that are documented in CASA Docs. However, we will also add a link to the CASA website, so that the memo series can be accessed directly from there as well. In addition to the memo series, we also started a Knowledge Base, where we gather useful information (e.g., based on development efforts or helpdesk tickets), which is not sufficiently detailed that it is suitable as a memo. This Knowledge Base will appear below the Memo Series starting with CASA Docs 5.7.

4. The Newsletter is well received by the CUC and should be continued in the current format. Readership statistics were still unclear. To be sure that the efforts are put to good use, a proper assessment of the readership base and time spent will be helpful, though that in itself will also cost time.

We will continue the Newsletter in its current format. For the readership statistics, we rely on Google Analytics. As we have shown at the past two CUC meetings, the statistics based on Google Analytics have its flaws. Unfortunately, we see no easy way around this. We doubt that it is worth the effort and investment to try and find better ways to obtain reader statistics. We note that not all external monitoring software may comply with NRAO regulations.

5. Archiving and indexing newsletters and memos on a searchable webpage could help users to search for specific topics. For example, the release and installation instructions for CASA6 are, at present, only to be found in the May 2019 Newsletter.

Since the first official release of CASA 6.0 in Dec 2019, the CASA 6 installation

instructions have been accessible and searchable on CASA Docs. For searches in CASA Docs, we implemented a handy search tool that can find specific topics per version of CASA. The CASA Memos are part of CASA Docs, and should be searchable the same way. The CASA Newsletter is archived on the CASA website. We are also hoping to improve searches done through Google.

6.3 CARTA

2. Could the v1.2 or v1.3 (if possible) release be demonstrated at the January 2020 AAS meeting? And/or at the EAS2020 in Leiden with help from Allegro?

CARTA v.1.2 was presented as a demo at the Jan 2020 AAS by Juergen Ott, who also serves as the NRAO representative on the CARTA Science Committee. Feedback was positive. We will stay in close communication with the CARTA team regarding future possibilities for demo's at international meetings.

3. The CUC recommends including a few practical functions such as 1D/2D Gaussian fitting in the spectral and image domains. It would also help attract users if CARTA can export PDF/FITS images or FITS/ASCII spectra to the client side.

The ideas and recommendations will be reconciled with the CARTA forward plan and incorporated in to the future roadmap.

6.4 Transition from CASA 5.x to CASA 6.x

2. The CUC suggests the development of a simple 'finding chart' to allow users to match their requirements to the current and future CASA versions.

CASA tool and task naming and contents have recently converged much more closely between the versions than was originally expected. Consequently, every commonly used function should be identical in name and usage between 5.x and 6.x. Where small deviations remain, we will be sure to list the mapping in the release notes.

Outside of tools and tasks, there may be isolated instances of users reaching in to CASA and making use of other classes that have now been marked as private. These methods are still accessible in the private area, but this designation should serve as notice that these are not treated as the official external interface of CASA.

6.5 Specific functionality: polarization and imaging

Polarization

2. Connecting with developers outside the NRAO-CASA team is imperative to make use of new ideas and momentum generated in, for example, the EHT collaboration. This functionality can benefit from the easy addition of third party software a dedicated CASA6 module. Leaving this under the responsibility of third party developers (like the VLBI tasks) will also remove maintenance issues.

CASA 6 offers the possibility to include a module for contributions to the CASA package from any user or organization. Such an external contributions module will not conform to the standard development, documentation, and test process of CASA, but will allow everyone to contribute in a meaningful way. If there are certain outsized contributions from a specific group (e.g. VLBI) it may make sense in the future to move those to their own modules.

Imaging

1. Parallelization of the imaging process will definitely help to decrease the runtime of ALMA and VLA pipelines. It is not clear if alternative imaging algorithms from outside NRAO (e.g. wsclean) will also be exploited to further optimize the pipelines.

There are not currently any plans to incorporate alternative imaging algorithms in to CASA or Pipeline processing in the near to mid-term. However, many exciting options continue to be developed that allow for new and interesting ways to deal with performance and scalability. CASA remains open to evaluating new methods in conjunction with the Algorithm Research and Development group and any stated stakeholder requirements.

2. During the meeting there was some discussion on whether tclean can be safely killed by the user. A fairly simple but significant improvement would be to clearly state this information in the documentation and on-screen output.

Tclean is generally safe to kill at any time with the possible exception of the brief instances during which data, in particular the model, are being written back to the MS. These brief instances happen in a way that would make screen output of warning information both noisy and awkward. The CASA team will investigate potential alternatives of capturing the kill interrupt and allowing data writes to complete before stopping the process, or separating the model write step from the rest of tclean so that it is always clear when the model is being written. This would make tclean always safe to kill. For the next CASA 5.7/6.1 release, we improved the overall tclean description page in CASA Docs and also included the above information.

3. An impressive amount of development is ongoing in the ARDG, and the work connects well to what was presented last year. Though some timelines have slipped there is clear progress. The presentation was quite condensed, and given the importance of ARDG for new functionality, this group might benefit from a presentation separate from imaging and verification testing. For example, it is not clear to the CUC what the status is of the 2018 “planned deliveries”, aw-projection and hybrid-Mueller imaging.

The ARDG group has no objection to giving a separate presentation at the next CUC face-to-face meeting. The ARDG roadmap document is being updated, and it will include a section on progress of the planned items in the last version. Note that the ARDG is outside the scope of the CUC’s mandate, but we are happy to provide this information as an FYI.

4. The VLASS imaging required w-projection, which was not initially foreseen. This raises the question how regular users should assess the need for w-projection? Is there a simple test they can run prior to imaging?

Decisions on VLASS imaging algorithms came after extensive observation, experimental data processing, and evaluation of results, using familiar data reduction procedures. The need for w-term correction and the number of w-bins required were determined by measuring source position shifts at the chosen pixel resolution for a source several arcmin away from the phase-center. There was not a simple test developed and choices ultimately depend on the situation and demands on output image characteristics. These topics are explored in publications, memos and discussion within the CASA documentation as well as regular data reduction workshops hosted by NRAO.

6.6 Archival Data Processing, Pipelines, and SRDP

1. Maintaining the capability to access old archival data is extremely important, especially with the increased attention that transient astronomy has been getting, and the fact that many of the old data have never been published. In particular, as older ALMA data can presently only be calibrated with older versions of CASA, which cannot be easily installed onto more recent operating systems, public users cannot easily use those older ALMA data. It is important to be able to reprocess old data with the latest CASA, as well as to recreate the original processing and results. The priority should be the former. This will also ensure that when the older ALMA observations (e.g., from Cycle-0,1) are used for complementing short-spacing information of the later observations, they can be jointly imaged with a consistent weighting scheme.

The CASA team realizes the need for users to re-produce archival data with new CASA versions, and will work with the Pipeline and SRDP teams to address this CUC

recommendation. We stress, however, that older ALMA data can be calibrated with any recent version of CASA. The problem is that in order to re-process older ALMA data with the standard archival scripts that are supplied by the ALMA observatory, users need to use specific older pipelines, which are shipped only with older CASA versions.

2. The CASA team should work with the ALMA observatory to identify issues preventing the re-reduction of archival data, including political ones, and develop solutions for enabling continued usefulness of the ALMA archive data.

We will raise this topic with ALMA Science Operations.

3. Ultimately, having the pipeline handle self-calibration in addition to the regular calibration would be desirable, but that is an ambitious goal. A first step might be for the pipeline to assess whether there is sufficient signal-to-noise for self-calibration and/or whether the target image is likely dynamic-range limited, and then print out a message to the user suggesting if self-calibration may be beneficial.

We will pass these comments on to the ALMA pipeline team. The ALMA pipeline is not within the scope of CASA or the mandate of the CUC. However, the pipeline team is beginning to plan what self-calibration might entail and we will learn more on the specific proposed solution in the coming year.

4. The balance of the development between pipeline and single-user operation requirements should be maintained.

We will continue to balance the needs and priorities of CASA's diverse group of stakeholders. The importance of maintaining CASA as a capable end-user data reduction package is understood.

5. It would be instructive to provide some brief introduction and/or explanation of the pipeline scripts and pipeline tasks executed together with the data delivered. The pipeline webpage should be made more known and public to general CASA users.

We will pass these comments on to the ALMA pipeline team. The pipeline team is in the process of developing and clarifying their documentation and website locations. We will provide links to relevant pipeline documentation in CASA Docs.

6. For the time being, the ALMA science archive should indicate "which version of CASA pipeline one needs to download" when fetching a specific data set.

There is a "Pipeline Version Tracker" available from the Science Portal at <http://almascience.org/processing/science-pipeline#version>, which lists the versions of CASA+Pipeline used in ALMA Operations, as well as the versions which should be used for

any restoring or reprocessing of the data from the same cycle. The Pipeline Working Group and Archive Subsystem Scientist are ultimately responsible for determining how to communicate pipeline compatibility to archive users. The CASA team attempts to satisfy backwards compatibility requirements for older data up to a reasonable point but there may be broader implications at the pipeline processing level that are outside the purview of the CASA team. We will pass along this recommendation to the relevant parties.

6.7 Documentation: CASAdocs and CASA Guides

1. We applaud the efforts to fully populate the online CASAdocs, and would like to see it extended to the rest of CASA. CASAdocs and CASA Guides are very helpful to users, and the tutorials are a very good introduction for new users/students.

We will continue to prioritize accurate documentation of CASA for users.

2. We recommend including more scientifically-oriented examples of the usage of the tasks, either in CASAdocs or CASA Guides. The present contents of the ALMA Guides are not sufficient for scientific purposes. Both simple and typical examples of the usage are necessary.

We are working towards CASA Docs being the complete description of the CASA code, which specifies for users how CASA works, and against which we will develop and test CASA code. CASA Docs is meant to be instrument-independent. As such, we made a conscious decision not to include telescope-specific examples on data reduction strategies in CASA Docs. Those examples are provided by the instrument teams, mostly in the CASA Guides. Also here, a conscious decision was made last year by the instrument teams to decrease the number of CASA Guides in order to improve clarity and accuracy. We believe that the examples currently provided in the CASA Guides *are* sufficient for most scientific purposes, and add that the NRAO and the Helpdesk have the resources in place to help users for which this would not be sufficient. However, we welcome specific feedback from the CUC on critical content that the CASA Guides are currently lacking, and will work with the instrument teams to add that to the CASA Guides.

3. It would be desirable to have a “search” in the CASA Guides (for example, search for “automasking”). The CASA Guides would also benefit from an index, a master list, and some links to and from CASAdocs.

We will discuss this recommendation with the instrument teams that maintain the CASA Guides, and explore what is possible in terms of search options within the existing wiki structure. In most CASA Guides, links to CASA Docs are included with the task names, but we will go over the Guides and include additional links to CASA Docs where relevant.

4. We recommend including documentation of NRAO pipelines. At least a sort of flow-chart of which tasks are executed in which order by the pipeline would be useful for users to understand what the pipeline is doing and assess what (if any) kind of intervention in the pipeline might be useful.

An improvement in documentation, including the description of pipeline tasks in CASA Docs, is in progress by the pipeline team.

5. Ideally, intuitive operations with CASA should not lead to erroneous results. If they do, it has to be very obviously cautioned/documentated. As a particular example, the ongoing issues with the virtual model column and tclean should be resolved. There should be an easy way for users to ascertain from an ms whether the virtual or the actual model column is populated, and what exactly it is that has been put into the model column. The tclean messages should be more explicit about whether the model has in fact been saved by the end of tclean.

The CASA team strives for intuitive behavior, but sometimes the result is a mix of specific needs at the time of development and constraints of the current implementation. We try to mitigate instances like these whenever possible as they are identified. Over the past few years the CASA team has placed tighter controls on new features and modifications and is working through technical debt as time and resources permit. For tclean in particular, we are evaluating options such as explicitly separating the tclean and predictmodel steps, and/or to make the relevant lines of text in the tclean logger more obvious than they currently appear. However, ideas such as this will need to be prioritized in our next development cycle.

6.8 CASA Reliability & Performance

1. The CUC is impressed with the new diagnostic and performance plots, this is a great improvement over last year. As a next step we would like to see a selection of the plots being shared with different user bases. Also, the tools for generating pipeline performance plots could be shared with communities who develop their own CASA based pipeline, e.g. via the new Memo series.

The CASA team agrees that the next logical step is to automate the generation of performance plots on a regular basis and publish them to users. We are in the process of moving in that direction. The pipeline team is also working on revamping their own documentation and automated test process, and we will discuss with them the potential to advertise similar performance reports at their level.

One component of CASA 6 is a casatests module that will allow users to execute verification

tests. Over time we hope to include more performance tests on the individual task level in this module.

2. The CUC is glad to see several new levels of functional tests. The testing team is expanding and actively developing their testing schemes. The addition of stakeholder tests is expected to help improve reliability. A minor concern is ensuring efficient communication among geographically widely distributed testing teams. Good communication and management tools are important to make this work and link to the development teams. For now, this seems to be well under control.

The globally distributed nature of the CASA team can be a communication and coordination challenge, but we strive to utilize the best tools and still retain frequent face to face contact whenever possible. This year, the Test Lead and one of the main test engineers met face-to-face at ESO. Test team members and leadership are expected to see each other face to face later this year.

3. The CASA telemetry reporting has now been active for over a year (since coming online shortly before the CUC meeting 2018). The CUC are pleased to see these results and note that the telemetry is providing a deeper insight into CASA usage and user behaviour than would be possible from only helpdesk feedback and user surveys. As noted in the 2018 CUC recommendations, telemetry provides a huge amount of data and the CUC will be interested to see how the data has been used to influence CASA development priorities.

At a minimum, the data collection provides justification for continued funding and application of resources towards the project. We are starting to influence the higher level development strategy as well, with one example being the modularization axis and clear usage pattern where many people seem to be using parts of CASA in isolation. What is more visible to users and the CUC is generally specific near term decisions, and these are likely to be the last to directly benefit from telemetry collection.

4. To compliment the telemetry reporting, the CASA crash reporting functionality is also providing useful information to the developers allowing the focussed effort in improving problem areas. The CUC note that from the data shown in the CUC meeting it is important that the CASA team ensure all crashes are being reported to the system. Few crashes involving the viewer were reported, which from general user experience seems incorrect. The CUC asks if the crash reporter is capturing crashes where a user 'kills' a task which has failed before a crash can be reported?

We agree that the crash reporter numbers may seem suspiciously low and we will investigate to find an explanation.

6.9 Support for Other telescopes

1. VLBI can also serve as an example for how external contributions can be merged into the CASA framework. There are concerns about the lack of use of CASA for VLBA. Further development of VLBI functionality can benefit from VLBA involvement (and vice versa).

VLBA staff have recently stepped up involvement in CASA planning and development by taking part in CASA stakeholder discussions and participating in monthly CASA/VLBI teleconferences with our JIVE development colleagues. VLBI work is being incorporated into the CASA development cycles and plans are in place to analyze and better understand the gaps that hinder broader VLBA adoption of CASA. During the planning for every fiscal year's budget we ask if it is time to add additional VLBA support for developments in CASA. We will discuss this next in late April.

2. There should be a clearer pathway for instrument-dependent parameters to be added to CASA for larger observatories such as MeerKAT, LOFAR, ASKAP etc., in particular primary beam models and calibrator sources that are not already included.

In the near term, the CASA User Liaison can serve as the communication hub for other instruments requesting updates to CASA. Longer term, the CASA team is looking at ways to standardize the definition of instrument and calibrator models with an interface for reading them in to CASA. This may begin as an SRDP initiative for NRAO instruments before expanding. However, a timeline is not yet in place.

3. The CUC wondered whether CASA team should reach out to single-dish observatories such as Green Bank or Parkes about using CASA for their data reduction.

The NRAO Assistant Director for Data Management and Software (Glendenning) will raise this topic with the Green Bank Observatory Director. In general we are not trying to “sell” CASA to observatories, although we are open to discussing collaborations with any observatories that would consider it to be in their self-interest.

6.10 CASA Next Generation Infrastructure (CNGI)

3. Discussions with external developers should have already started. They can provide useful feedback on potential issues and benefits of choices that are now made. This may include discussion of MSv3.

Discussions with external development teams have indeed begun, including the MS/casacore

consortium and the astropy community. Feedback has been very positive with a strong desire for collaboration and enthusiasm for the preliminary design choices. We intend to hold an international meeting in 2020 to discuss the mechanisms of collaboration with potential partners after CNGI has passed an internal design review at the NRAO (i.e., after we know we are ready to proceed).

6.11 CUC Membership Term

"The CUC also discussed extending the duration of service for members. If CUC members are serving 4-year terms or, alternatively, serving up to an additional three years, then it will maintain experience on the committee while still allowing for new members to cycle on. The CUC recommends that the CASA team extend the duration of terms for CUC members to an optional fourth year."

The NRAO Assistant Director for Data Management and Software will draft an update to the terms of reference with this recommendation, and discuss it with the ALMA/NRAO science operations personnel who appoint the CUC members.

7 Recommended Priorities for the coming year

1. Implementation of CARTA
2. Ensuring Access to Archival Data in CASA
3. A smooth transition to CASA 6
4. Continued improvement of documentation and communication of CASA capabilities and reliability

We thank the CUC for summarizing the main recommended priorities for the coming year. We will continue to work closely with the CARTA and Pipeline teams to address points 1 and 2, and will put emphasis also on addressing points 3 and 4.