



<b>Title:</b> Algorithms R & D Group Management Plan	<b>Author:</b> S. Bhatnagar	<b>Date:</b> 01/23/2018
NRAO Doc. #:		<b>Version:</b> Final

# Algorithms R&D Group Management Plan

<b>PREPARED BY</b>	<b>ORGANIZATION</b>	<b>DATE</b>
Sanjay Bhatnagar	NRAO, Socorro	01/23/2018

<b>APPROVALS (Name and Signature)</b>	<b>ORGANIZATION</b>	<b>DATE</b>
Brian Glendenning	NRAO DMS	01/23/2018

**TABLE OF CONTENTS**

**1 Overview ..... 3**

**2 Purpose and Scope ..... 3**

**2.1 Deliverables ..... 3**

**3 Group organization..... 4**

**4 Management Plan ..... 4**

**4.1 Development of Scientific and Technical Priorities ..... 5**

**5 Technical Plan..... 5**

**Appendix..... 7**

# I OVERVIEW

As the name suggests, the Algorithms R&D Group (ARDG) at NRAO was formed to address requirements driven by NRAO mission and goals which also need research and associated development for data processing algorithms. The group consists of NRAO scientists with skills and research experience in algorithm development and deployment at the observatory. The ARDG has been operational for over a decade at NRAO, initially as an informal group, and as a sub-group under the CASA management for the past few years.

The scope, time-lines and priorities of the work done by the ARDG is significantly different from that of the production software group. The type of personnel, in terms of their skill-set required for ARDG to be effective is therefore also quite different. It also requires a different management and reporting structure. Making ARDG a distinct group separate from CASA within the DMS is therefore necessary. This document lays out the management plan for the new ARDG.

## 2 PURPOSE AND SCOPE

The primary purpose of the ARDG is to develop and execute a long-term plan for algorithms R&D based on NRAO long term priorities and mission. Such high-level priorities however often lead to only very generic end-goal (e.g., algorithm/software that would deliver imaging performance consistent with the thermal noise limit of the telescope) and it requires significant R&D effort and longer-term focus to arrive at a practically useful and usable solution with intermediate check-points and deliverables. In this sense, the ARDG is distinct from production software groups like the CASA group, which is tasked with delivering the algorithms developed by the ARDG and others to the users via a stable software product. CASA is therefore driven by shorter-term priorities and deliverables. On the other hand, algorithms R&D depends on the CASA group for the software infrastructure and existing and proven algorithms for carrying out further research for new algorithms. ARDG is therefore a customer of CASA for the research phase and CASA is a customer of the ARDG for the development and deployment phase of algorithms. Delivering a complete description of new algorithms shown to adequately solve the problems they address, including proto-type code, documentation and help where necessary for deployment, is therefore the primary deliverable of the ARDG.

The scope of the work for the ARDG would directly include work on calibration and imaging algorithms, RFI flagging and/or mitigation, related computing technology investigations, development of prototypes and recommendations, and possibly image analysis particularly where feedback into calibration, imaging or computing technology is expected or foreseen. The scope also includes investigation of techniques and algorithms for efficient implementation, e.g. on massively parallel computing platforms like the modern GPUs and multi-node CPUs. Apart from the EVLA and the ALMA telescopes, the scope of work will also include next generation telescopes like the ngVLA, upgrades to ALMA or EVLA, which may require algorithmic developments for success, or other telescopes in which NRAO invests resources.

### 2.1 Deliverables

The ARDG deliverables would include

- Publications of the results, preferable in a refereed scientific journal in a timely fashion, with a description and demonstration of a working algorithm.
- A working prototype code based on CASA code-base, including regression tests which can be run by a competent third party (e.g. members of the CASA group) and test data – both before and

after processing – and scientific products to help verify correctness of an implementation.

- A memo with details of the software interface, and
- Any commissioning assistance required.

Additionally, for algorithms that require HPC, similar products will be developed for deployment on HPC platforms.

ARDG Lead and/or a group member would be responsible for representing NRAO at various meetings, both scientific and administrative. ARDG discussion forum will also be revived for discussions related to post-processing algorithms and related computing and software needs. This will provide individual scientists/smaller groups a forum to discuss the problems/issues they face with ARDG scientists who may be able to address them.

### **3 GROUP ORGANIZATION**

The ARDG currently is, and expected to remain a relatively small group consisting of 4-6 scientific staff actively involved in both algorithmic and observational astronomy research. The group organization will be flat with minimum hierarchy necessary for timely development of long-term plans in accordance with NRAO priorities and goals, reporting and periodic tracking of the plan execution and any minor changes that may be required.

The ARDG will be under the DMSD currently headed by B. Glendenning. The ARDG will have a Group Lead (currently, S. Bhatnagar) who will report directly to the DMSD AD. Members of the group will typically have 50% of their functional time assigned to the ARDG and they will report to the ARDG Lead for this fraction of their functional time. The current ARDG staffing details are in the table below.

<b>Name</b>	<b>Percentage of functional time</b>
S. Bhatnagar (Scientist)	50%
U. Rau (Associate Scientist)	50%
K. Golap (Scientist)	50%
P. Jagannathan (Assistant Scientist, new hire)	100%
B. Kirk (SE-I, 1-year term position)	100%

The ARDG Lead will preferably be also an active researcher himself/herself to ensure that the group is effective locally at NRAO and competitive globally. The group Lead will be responsible for maintaining a good scientific and technical overview of the group's activities and priorities and be able to provide a general direction for research by the members of the group.

### **4 MANAGEMENT PLAN**

The ARDG Lead will hold regular group meetings, both to ensure communication with-in the group and keeping herself/himself up to date with the activities of the group. The Lead will also hold quarterly meetings with the AD, DMSD to transfer the information about the group's activities and progress to the upper management. These management activities are expected to cost 20-25% of the function time.

The group Lead will be responsible for developing and maintaining a five-year road map for the group and track its execution. This plan will be reviewed on a yearly basis in collaboration with the DMSD and SSR

ADs, CASA Lead and few relevant members of the scientific staff. Any major change in this plan may be discussed and reviewed as needed.

Details of the actual research and decisions will be made by the individual members in consultation and collaboration with other members of the group as well as with other resources outside the group. In this, the ARDG Lead will act like any other member of the group as well as a scribe for recording any useful details of the activities and required change in the overall plan necessitated due to new developments and resulting understanding of the subject matter.

The development work, even in the prototype stage, done by the members of the group will be done using the CASA code-base. This is to minimize (a) development overheads for the research activities, and (b) cost of transferring the information and code to the CASA group once an algorithm matures for incorporation and maintenance in the production software. Experience shows that this has significant overlap and strong dependence on the decisions, direction and short-term activities of the CASA group itself. The ARDG as a group will therefore maintain close co-ordination with the CASA group and when deemed more efficient, may contribute directly to the CASA production branch for short periods. While member-to-member contact between the ARDG and CASA Group will probably remain the most efficient mechanism, the ARDG Lead will co-ordinate with the CASA group for coherent development on the ARDG.

#### **4.1 Development of Scientific and Technical Priorities**

The ARDG activities are forward-looking with an attempt to foresee scientific and technical needs of the observatory sufficiently before they become operation blockers or significant impediment in effective and efficient use of NRAO facilities. The development of a 5-year road-map for the ARDG therefore requires input from members of the scientific staff and external users involved in projects that push some existing boundary (scientific, technical or computational), the CASA Group, NRAO HPC Group, the DMS and SSR management and the members of the ARDG itself. Post-processing in general and algorithms R&D for it in particular is relatively specialized area requiring multi-disciplinary skills ranging from mathematics, astronomy and computing to software engineering.

With a large scientific staff composed of members with very diverse skill-set and experience levels, it is impractical to involve the entire scientific staff in developing this forward-looking road map itself. Therefore, in consultation with the SSR and DMS leads, we will work with representative staff scientists for developing the road map itself. An Advisory Committee (AC) will be appointed by AD, DMSD with representatives from these entities. The ARDG Lead will be responsible for developing a five-year road map with the scientific and technical priorities for the group, taking into consideration the inputs from the AC and the ARDG members. Since the ARDG workflow is longer-term in nature, attempt will be made to include reasonable comments/suggestions from the larger scientific staff during the R&D phase.

The ARDG Lead will also be responsible for maintaining this road map with a yearly review of it involving DMS and SSR divisions, with updates on a shorter timescales as necessary. This will include briefing the NRAO scientific staff at least annually and presentations to internal or external review committees as needed on the progress against the road map and any other related updates.

### **5 TECHNICAL PLAN**

The ARDG work will be done using the CASA code base as far as possible. This is to ensure that the prototype code is as close to the production software as possible and to benefit from a wide range of algorithms and infrastructure that exists in CASA during R&D stage. This however also requires that the

ARDG will become a customer of CASA and a stakeholder in the plan and prioritization process for the CASA group. On a day-to-day optional level, this would also mean that development in both ARDG and CASA is done in close co-ordination between the two groups. The DMSD is the parent division for both these groups and it will be responsibility of the DMSD management to facilitate this co-ordination.

ARDG development typically requires a longer gestation period before the prototype code is even usable. Since the work often is also product of original research, there may be a need to keep access to the code private (within NRAO). Since CASA now uses Git as the revision control system, ARDG will also use Git and may maintain a long-running branch. The branch will be managed by ARDG. Once the research-phase is considered finished (typically after scientific publications describing the algorithm and demonstrating its scientific application), the code may be made public. When released in CASA the code will be public; before this point, its visibility will be a discretion of the ARDG Lead.

The CASA stakeholders will define the relative priority of getting new ARDG functionality into CASA. This, by design, is expected to be high priority (since the ARDG road map will include input from CASA stakeholders). The CASA group is responsible for converting the research code-base to a production code (possibly with ARDG scientist's efforts, but on CASA time), maintaining it and providing user support via established mechanisms for these in CASA.

For projects where there is no need to maintain a separate private branch, the code will be in the CASA repository and we do not anticipate extra help from the CASA build and distribution team. For branches that may be private and separate from the CASA repository, some help may be necessary from the CASA systems team. We anticipate only light-weight periodic help here.

As in the past, ARDG personnel may periodically require purchase of higher-end computing hardware relative to those available for day-to-day processing at NRAO, or access to such existing hardware. This has so far not been a problem and combination of SSR and DMS division has been able to fund such needs. We expect this would continue – though the nature of the required hardware may change (e.g., from machines at NRAO to purchasing computing time at AWS). This may also be in the form of needing larger compute cluster resources (for testing and verification) than is typically allowed. Past experience suggests that this access is necessary for R&D and often also provides useful feed-back both to the HPC and CASA groups. Again, this has not been a major problem so far and we anticipate the same in the future. Any ARDG special needs in excess of \$20K will be defined at the start of every budget cycle.

Finally, proper testing often requires access to telescope test time to acquire test data, particularly in the later stages of development. This also so far this has not been a problem with the EVLA. This has not yet been tried with ALMA where the process of acquiring test-data is more involved.

Verification testing in the R&D stage of the work using the prototype code will be the responsibility of the ARDG. As now, any testing for validation of the code after hand-over to the CASA group will remain the responsibility of the CASA group. In addition to the verification tests developed by the ARDG, the ARDG staff will also provide any necessary and reasonable help in designing validation tests.

# APPENDIX

## RACI Matrix

**R: Responsible**   **A: Accountable**   **C: Consult**   **I: Inform**

Activity	ARDG			SSR			DMS		
	Lead	Staff	Adv. Cmte.	AD	External Users	Sci. Staff	AD	S/W DH	CASA Lead
Develop and maintain a long-term road-map	A,R	C	C	I	I	I	C	I	I
Appoint Advisory Counsel	C	I	I	C			AR	I	C
Track progress	AR	I	I	I	I	I	I	I	I
Organize yearly review of the road-map and any changes	R	I	C	I		I	A	I	C
Publish status and plans	AR						C		
Presentations within and outside NRAO	AR	C		C			C		
Co-ordinate with the CASA Group	R	C	I				A	R	R
Organize scientific and technical discussion forums	AR	C	C				I	I	I
Scientific and technical R&D work	A	R							
Scientific testing and verification of algorithms	A	R		I			I	I	I
Publish algorithms, develop prototype code	A	R	I	I	I	I	I	I	I
Transfer to CASA	R	R	I	I	I	I	C	C	AR
Maintain production code, user support, documentation	C	C		I			I	I	AR