Atacama Large Millimeter Array	ALMA-SW-xxx Revision:4.0 2002-03-06 *Requirements* Steven T. Myers

DRAFT — ALMA Offline Data Processing Requirements

Requirements

ALMA Pipeline and Offline Software Group:
B. Clark, F. Gueth, M. Momose, S. Myers,
P. Schilke, K. Tatematsu
ALMA Software Science Requirements Committee:
H. Ezawa, T. Handa, R. Lucas, J. Mangum, T. Ohnishi,
J. Richer, S. Scott, F. Viallefond, M. Wright
ALMA Software Use Case Group:
G. Harris, D. Muders, J. Schwarz, R. Warmels

Keywords: Requirements, Offline, Software, Science, Calibration

Author Signature: Steven T. Myers	Date: 2002-03-06
Approved by: B. Glendenning, G. Raffi, K.-I. Morita	Signature:
Institute: NRAO, ESO, NAOJ	Date:
Released by:	Signature:
Institute:	Date:

Change Record

Revision	Date	Author	Section/ Page affected	Remarks
1	2001-03-06	S.T. Myers	All	initial version
1.2	2001-05-01	S.T. Myers	All	version for SSR telecon
2	2001-06-21	S.T. Myers	All	version to SSR
2.2	2001-07-12	S.T. Myers	All	comments from SSR
3	2001-07-24	S.T. Myers	All	after Berkeley, remove pipeline
3.1	2001-08-08	S.T. Myers	All	finish first-round edits
3.2	2001-08-31	S.T. Myers	All	incorporate comments
3.3	2001-09-02	S.T. Myers	All	regularize phrasing, for ASAC
3.4	2001-12-06	S.T. Myers	All	comments, for review
4.0	2002-03-06	S.T. Myers	All	after review, incorporate comments

`$Id: myers-report-4.0.tex,2002/03/06 myers Exp myers $`

Contents

1 Introduction **5**

1.1 Nomenclature . . . 5

1.2 General Considerations . . . 6

2 Offline Data Processing Requirements **8**

2.1 General Requirements and Interaction with Other ALMA Elements . . . 8

2.1.1 Goals of the Offline Package and Relation to the Pipeline . . . 8

2.1.2 Operational Issues . . . 8

2.2 Interface . . . 11

2.2.1 General User Interface Requirements . . . 11

2.2.2 Graphical User Interface (GUI) . . . 11

2.2.3 Command Line Interface (CLI) . . . 12

2.2.4 Interface Programming, Parameter Passing and Feedback . . . 13

2.2.5 Documentation and Help Facility . . . 13

2.3 Data Handling . . . 15

2.3.1 General Data Requirements . . . 15

2.3.2 Data Import and Export . . . 17

2.3.3 Images and Other Data Products . . . 18

2.3.4 Foreign Data . . . 18

2.3.5 Interaction with the Archive . . . 19

2.4 Calibration and Editing . . . 20

2.4.1 General Calibration and Editing Requirements . . . 20

2.4.2 Atmospheric Calibration . . . 22

2.4.3 Interferometer Data . . . 22

2.4.4 Single Dish Data . . . 23

2.4.5 Mosaicing Considerations . . . 24

2.4.6 Ancillary and Diagnostic Data . . . 24

2.5 Imaging . . . 25

2.5.1 General Imaging Requirements . . . 25

2.5.2 Interferometer Imaging . . . 26

2.5.3 Mosaicing and Single Dish Imaging Considerations . . . 26

2.6 Data Analysis . . . 28

2.6.1 General Analysis Requirements . . . 28

2.6.2 Spectral Line Analysis . . . 28

2.6.3 Image Cube Analysis and Manipulation 29
2.7 Visualization 33
2.7.1 General Visualization and Plotting Requirements 33
2.7.2 Display Appearance and Interactivity 33
2.7.3 Visibility Data 35
2.7.4 Other ALMA Data 35
2.7.5 Image-cube Manipulation 36
2.8 Special Features 37
2.8.1 Simulation 37
2.8.2 VLBI 37
2.8.3 Solar System Object Observing 37
2.8.4 Pulsar Observing 38

1 Introduction

The history of this document and correspondence can be found at:

`http://www.aoc.nrao.edu/ smyers/alma/offline-req/`

This document is `myers-report-4.0.tex`.

This document describes the requirements for offline data processing software packages in order to be able to handle the ALMA data output. It is assumed that there will be (at least) one software package available for users to reduce their own data and/or data from the archive offline at their home institution, at an ALMA regional center, or remotely using ALMA center computing. This suite of tools must fulfill the Offline Data Processing Requirements (Section 2). We refer to the ensemble that fulfills these as the "Package". Again this might be an assortment of different programs from different software packages (e.g. AIPS, GILDAS, MIRIAD, AIPS++) but in this case it is highly desirable that there be at least one single package that fulfills the Offline requirements or that there be an installation that integrated the necessary parts. Note, however, that is is highly unlikely that disparate applications from different packages will fulfill the requirements on similar "look and feel" and inter-connectivity, and thus it is likely as well as desirable that this will be a single homogeneous suite.

This document assumes requirements already delineated in ALMA-SW MEMO 11 "ALMA Software Science Requirements and Use Cases" at:

`http://www.alma.nrao.edu/development/computing/docs/joint/0011/ssranduc.pdf`

Much of the content of this document is based on the AIPS++ User specifications Memo 115 found at:

`http://aips2.nrao.edu/stable/docs/specs/specs.html`

Although it was intended to be the basis for AIPS++, these are an excellent starting point for our requirements to build upon. We do not think this makes this too specific to AIPS++, as our document is intended to be package-independent.

1.1 Nomenclature

The subject of Section 2 on Offline Data Processing Requirements is referred to as the *Package* or *Offline Package*. This is intended as a set of tools or programs, believed adequate for ALMA reductions, and used by observers for science and by ALMA staff for reductions upon which the behavior of the system will be judged. It may consist of packages provided by different groups, with transitions provided to integrate them into a single suite. It may be that more than one Package fulfills the requirements of Section 2 and thus can be considered as suitable for ALMA processing. The requirements will state that the Package will be available for installation on the observer's own computer systems as well as present at ALMA centers.

There is another set of requirements, given in ALMA-SW MEMO 11, for the software that forms the ALMA Online Data Processing Pipeline, or simply the *Pipeline*. Note that the Pipeline may or may not be based on the offline Package(s), depending on implementation. However, the functionality of the Pipeline must be available in the offline Package for users.

The user of the Package may be referred to as *user* or *observer*, and may be the actual proposer or a staff member or an archive user. In automated modes, the user may actually be another tool or program.

The *Archive* refers to the totality of the ALMA data storage and consists of possible different physical archives.

The definitions of various terms, such as *scans*, are given in ALMA-SW MEMO 11.

1.2 General Considerations

A. Two fundamentally new aspects of ALMA are the integrated Archive and the Pipeline, therefore the impact of requirements on these two areas should be considered. We have included topics "Relation to the Pipeline" (§2.1.1) and "Interaction with Archive" (§2.3.5).

B. There is a operational difference between running the ALMA array in interferometric and single-dish modes. However, this difference may not be so fundamental for many types of data processing. For example, if several single dishes observe together, much of the calibration will be done interferometrically (pointing, focus, beam shape); for interferometric observations, temperature scale is derived from single-dish measurements. Moreover, the inclusion of single-dish capability in ALMA is primarily intended to aid in the mosaicing of targets larger than a primary beam, and thus array and total-power data are likely to be combined in these instances. Therefore, these two paths are split or combined as is appropriate for the context, with special sections on "Mosaicing Considerations" also included to clarify these issues.

C. There is no fundamental difference between spectral-line and continuum observations, merely number of channels and bandwidths coming out of the correlator. Due to the nature of the ALMA correlator, most of the special calibration needs of traditional spectral line observations (e.g. bandpass calibration) are also applicable to continuum observations (continuum is built up through summation of spectral channels taken at low resolution). We will assume that all data will be effectively taken in spectral line mode.

D. There is no fundamental distinction needed for polarization data, merely consideration of the number of polarization products or polarization states needed for processing. The model we consider here is the processing of one or more Stokes parameters and thus polarization is integrated into all the topics. One complication that must be considered is the combination of the array (antenna polarization correlation products) and single-dish data (the antenna polarization outputs).

E. We use a system of prioritizing with codes:

1 = essential
2 = highly desirable
3 = desirable, but not critical

These codes are enclosed in brackets [] at the end of the items they qualify.

It is intended that Priority 1 items must be present in the Package and work with high efficiency. Priority 2 items should be in the Package, though there may have to be sacrifices in performance or availability may be delayed. Fulfilling Priority 2 items well should give a boost to a given Package's rating. We expect that the Package will fulfill all Priority 1 and 90% or more of Priority 2 requirements. Priority 3 items should be considered for upgrades or development.

F. The requirements in this document are prefixed with **OL**. Specific requirements are broken into subcases or instances (e.g. R1 R1.1, R1.2) for clarity. These sub-requirements are either special cases of the main requirement (often with different priority) or itemized instances or examples related to the requirement.

G. There are a number of instances where certain specifications, such as data formats supported, are designated as being in a list maintained by the ALMA project. There are other points in the requirements that refer to "standard modes", which implies a detailed list (tbd) also maintained by the project. These cases are indicated in **boldface** to aid in determining which lists will need to be constructed.

H. There are a number of places where we discuss the philosophy behind our choice of requirements. In those cases, that discussion is given as *italicized* text and is not meant as a requirement in and of themselves.

I. At the time this document was written, the ACA has not been officially scoped as part of ALMA. However, some placeholder ACA-related requirements are included in the main requirements section, and are clearly labeled as such.

2 Offline Data Processing Requirements

2.1 General Requirements and Interaction with Other ALMA Elements

2.1.1 Goals of the Offline Package and Relation to the Pipeline

Note: An ALMA Offline Data Processing Package (or "the Package") is primarily intended to enable end-users of ALMA (e.g. observers or Archive users) to produce scientifically viable results that involve ALMA data products. The secondary use is to enable ALMA staff to assess the state of the array and derive calibration parameters for the system, although it is anticipated that the Pipeline will be the primary engine for Online Data Processing and system monitoring.

Also note that not all offline analysis tools will necessarily be in the Pipeline Package. For example, one of the important differences between Pipeline and Offline reduction path is that Offline one should have extensive interactive capabilities to merge and compare data with different resolution, coordinate system, data grid, and so on.

OL–1.1–R1 There must be an Offline Data Processing Package that fulfills the requirements laid forth in this document.

Priority: 1

OL–1.1–R2 All standard observing modes supported by ALMA must be processable by the Package.

Priority: 1

OL–1.1–R3 The Package shall be installable at the users home institution and available at ALMA regional centers (both locally and remotely). It shall be portable to **supported platforms designated by the ALMA Project**, including systems without network connections and laptops.

Priority: 1

OL–1.1–R4 The performance of the Package shall be quantifiable and commensurate with the data processing requirements of ALMA output at a given time. This will be benchmarked (e.g. "AIPSmarks") and reproduce accurately results for a fiducial set of reduction tasks on specified test data.

Priority: 1

OL–1.1–R5 All standard processing functionality available in the Pipeline shall be available in the Package also as an offline analysis option.

Priority: 2

2.1.2 Operational Issues

OL–1.2–R1 Installation of the Package by an end user must be straightforward, preferably without special (e.g. root) user permission.

Priority: 1

OL–1.2–R2 Error reporting and handling shall be user-understandable and non-destructive at all levels in the Package.

- **OL–1.2–R2.1** Error reporting messages shall be written for end users, not programmers.
  Priority: 1
- **OL–1.2–R2.2** There must be provision for job control such as interrupt and abort.
  Priority: 1
- **OL–1.2–R2.3** Error handling shall be non-destructive (data shall not be left corrupted, and recovery of recent changes be available). Common failure modes (as enumerated: invalid application

parameters, exceeding of resource limits (disk/memory), and algorithm failure modes (e.g. no convergence)) should be handled gracefully.
Priority: 2

OL–1.2–R2.4 Code traceback of execution errors shall be available. This should be geared to the effective reporting of failure modes and bugs by the user.
Priority: 3

OL–1.2–R3 There shall be session logging, including the following features:

OL–1.2–R3.1 logging of commands and user inputs shall be provided
Priority: 1

OL–1.2–R3.2 logging of tool results such as success or error, files written, time of completion shall be provided
Priority: 1

OL–1.2–R3.3 logging of tool output such as summaries of results shall be provided
Priority: 2

OL–1.2–R3.4 the session log shall be readable by the user (ie. in a text file, not in a binary format)
Priority: 2

OL–1.2–R3.5 session logs shall be executable by the package UI, reproducing the entire session
Priority: 3

OL–1.2–R4 There shall be comprehensive handling of multiple users and multi-tasking with access and process control.

Priority: 1

OL–1.2–R5 The Package providers shall provide support for the Package, with bug-fixing on timescales appropriate to the severity level of the defect (e.g. 1 week or less for catastrophic bugs with no work-around). **These defect levels and timescales should be delineated by the ALMA project**.

Priority: 1

OL–1.2–R6 The Package providers shall provide timely improvements and updates based on user feedback. There shall be a path for the ALMA Project, for its own use and as a proxy for the users, to influence the development cycle of the Package. It is the responsibility of the Project to negotiate needed improvements with the Package providers.

Priority: 1

OL–1.2–R7 Backwards compatibility of core package components should not be broken without compelling scientific reasons. Tools should be provided to parse user scripts and warn of package changes.

Priority: 1

OL–1.2–R8 Source code for the astronomical routines in the Package shall be available to the user.

Priority: 1

Notes: This does allow for the use of some proprietary data handling, output formatting, and special processing (e.g. Pixon) routines.

OL–1.2–R9 User installation of the basic package shall not be restricted by other issues such as expensive or unduly restrictive licenses. The Package license should convey all other necessary licenses (such as GNU) or they should be available to the user for only a nominal fee.

Priority: 2

Comment: This allows for the use of commercial components. However, a balance must be struck between excessive costs to the project and requiring some cost to be borne by the user, should this issue arise.

OL–1.2–R10 There shall be the provision for the development and incorporation of user-supplied code.

Priority: 2

OL–1.2–R11 The application of successive stages of calibration, correction, flagging and editing shall not be destructive to the data. The Package should be able to recover and revert to earlier stages without resorting to saving of the entire dataset at intermediate states.

Priority: 2

2.2 Interface

2.2.1 General User Interface Requirements

OL–2.1–R1 User must be able to choose from a variety of interface styles, including:

- **OL–2.1–R1.1** A Command Line Interface (CLI) must be provided, with access via both an interactive input and via script.
  Priority: 1
- **OL–2.1–R1.2** A Graphical User Interface (GUI) must be provided for interactive processing. Actions taken under the GUI must be loggable, editable, and executable by the CLI.
  Priority: 1

OL–2.1–R2 The user shall be able to interact with the host operating system with command sequences invoked from the UI.

Priority: 1

OL–2.1–R3 Multitasking for all interfaces shall be available where appropriate. It must be possible to run one or more long-running calculations in the background. While background tasks are running normal interactive activities must be possible.

Priority: 1

2.2.2 Graphical User Interface (GUI)

Note: The GUI is intended to be the primary interactive mode for users, especially neophyte users, and thus should be tailored for clarity and ease of use. Use of the Package GUIs should be a pleasurable, not a frustrating, experience!

OL–2.2–R1 The use of the GUI shall not entail an excessive learning curve. Average users, with experience with the current generation of packages (e.g. AIPS, Miriad, GILDAS, IRAF) shall be able to become proficient in GUI use in a timescale of approximately 12 hours dedicated use, and truly neophyte users (e.g. graduate students) should be reach proficiency with an investment not exceeding 40 hours of dedicated use.

Priority: 1

OL–2.2–R2 The GUI shall provide real-time feedback via standard compact displays:

- **OL–2.2–R2.1** Window updating must be fast (less than 0.1s on same host).
  Priority: 1
- **OL–2.2–R2.2** Windows shall not take up excessive screen space, with full GUI controls visible on a window one-third the size of a standard view surface (or approximately 800×800 pixels).
  Priority: 1
- **OL–2.2–R2.3** Users shall have the choice of cascading windows or re-use of a single window for new operations.
  Priority: 2
- **OL–2.2–R2.4** Moving and resizing of all windows must be available, robust, and easy.
  Priority: 2
- **OL–2.2–R2.5** There shall be a master control GUI for process control which keeps track of subwindows and tools.
  Priority: 3

OL–2.2–R3 It must be easy to run GUIs remotely from the host machine (e.g. via X displays).

Priority: 1

OL–2.2–R4 GUI-based tools shall be available for reduction of data taken in all standard ALMA observing modes.

Priority: 2

OL–2.2–R5 The look and feel of the GUI will be uniform throughout the entire package.

Priority: 2

OL–2.2–R6 The look and feel of the GUI must be acceptable to both novice and more advanced users. The GUI mode might be customizable (perhaps through menu selection for "novice" or "advanced" mode) or a different simpler set of GUI tools might be available for beginners. The GUI features for beginner mode shall include:

OL–2.2–R6.1 built-in help facility with access to novice-oriented help documents (e.g. sections in the cookbook)
Priority: 2

OL–2.2–R6.2 sensible defaulting of values for parameters, with guidance for user choices where needed
Priority: 2

OL–2.2–R6.3 integrated functionality built around common data analysis tasks (e.g. single-field spectroscopic observations with fast switching, OTF mosaicing in continuum mode, snapshot observations of a large number of targets in single-field continuum mode)
Priority: 2

OL–2.2–R7 It is shall be easy for users to develop and include their own custom GUIs in the Package.

Priority: 3

2.2.3 Command Line Interface (CLI)

Note: the CLI is the primary mode for automatic reduction for ALMA, and it is anticipated that there will be a suite of "standard" scripts developed to help users in data reduction tasks. Thus it is important that the Package support all of its critical modes in the CLI.

OL–2.3–R1 The interface must have the facility to read in command files for batch processing of a sequence of CLI commands.

Priority: 1

OL–2.3–R2 The CLI shall have command-line recall and editing, with name completion where appropriate.

Priority: 1

OL–2.3–R3 All functionality of the GUI must also be available in CLI mode (although possibly with loss of simplicity in instances where the graphical selection is important).

Priority: 1

OL–2.3–R4 The CLI shall be usable remotely over low-speed (5400 baud) modem lines or network connections, with ASCII terminal emulation.

Priority: 2

OL–2.3–R5 A CLI mode menu for display and editing of input parameters and execution of tasks is desirable.

Priority: 3

2.2.4 Interface Programming, Parameter Passing and Feedback

OL–2.4–R1 The UI must have basic programming facilities such as:

- **OL–2.4–R1.1** variable assignment and evaluation
  Priority: 1
- **OL–2.4–R1.2** array handling
  Priority: 1
- **OL–2.4–R1.3** conditional statements
  Priority: 1
- **OL–2.4–R1.4** control loops
  Priority: 1
- **OL–2.4–R1.5** string manipulation
  Priority: 1
- **OL–2.4–R1.6** user-defined functions and procedures with argument or parameter passing
  Priority: 1
- **OL–2.4–R1.7** process control, interrupts, error handling
  Priority: 1
- **OL–2.4–R1.8** standard mathematical operations and functions
  Priority: 1
- **OL–2.4–R1.9** efficient special vector and matrix operations
  Priority: 2
- **OL–2.4–R1.10** user-defined data structures
  Priority: 2

OL–2.4–R2 Commands executed will be logged, with provision to re-execute the session.

Priority: 1

OL–2.4–R3 Input parameter and syntax checking shall be effected upon function calling or parsing with reporting of incorrect, suspicious or dangerous choices before execution where possible.

Priority: 1

OL–2.4–R4 Parameters shall be passable between applications in as transparent a manner as possible. However, global variables should not be the default, unless designated specifically by the user-programmer.

Priority: 1

OL–2.4–R5 Application variables shall be named consistently and as clearly as possible indicating their intended use using astronomical terms where appropriate.

Priority: 2

OL–2.4–R6 Parameter inputs to tools shall be stored for later recall.

- **OL–2.4–R6.1** Tool inputs shall be saved on closure, and reinstated on the next instantiation of the tool.
  Priority: 2
- **OL–2.4–R6.2** It shall be possible for the user to save the state of the parameters for the entire package, as well as for individual tools, as a named set (e.g. SAVE/GET in AIPS), and to recall these when desired.
  Priority: 3

2.2.5 Documentation and Help Facility

OL–2.5–R1 There shall be a variety of help levels and documentation formats accessible from the UI and over the Internet, applicable to novices, experts, and technical users. These shall include:

OL–2.5–R1.1 user cookbooks with extensive examples
Priority: 1

OL–2.5–R1.2 application descriptions and reference manual (with all inputs to functions and tools)
Priority: 1

OL–2.5–R1.3 online help, FAQ, email contacts
Priority: 1

OL–2.5–R1.4 release history, bug reports and tracking, patch descriptions
Priority: 1

OL–2.5–R1.5 programmer references and guides
Priority: 2

OL–2.5–R1.6 data format descriptions
Priority: 2

OL–2.5–R1.7 algorithm descriptions
Priority: 2

OL–2.5–R1.8 newsletters, email exploders, notes series
Priority: 3
Note: these would be maintained by the Package providers, with help from the ALMA project.

OL–2.5–R2 The Package creators must ensure the documentation is up-to-date and complete for all parts of the Package.

Priority: 1

OL–2.5–R3 Help materials shall also be available in printable formats, including

OL–2.5–R3.1 standard document formats (pdf, postscript)
Priority: 1

OL–2.5–R3.2 printer-friendly versions of HTML pages
Priority: 2

OL–2.5–R3.3 popular proprietary formats (MS-Word)
Priority: 3

OL–2.5–R4 Help shall be context-sensitive where relevant. In GUI mode, fly-over banners should indicate use of buttons and fields, and clickable help buttons should be available on all pages.

Priority: 1

OL–2.5–R5 In GUI mode, help functions may direct a browser to a Web page. In CLI mode, the Package must support in-line text based help also.

Priority: 1

OL–2.5–R6 Full search capability must be built into the documentation library.

Priority: 2

2.3 Data Handling

2.3.1 General Data Requirements

OL–3.1–R1 The Package must support data taken in any of the standard ALMA hardware modes.

Priority: 1

OL–3.1–R2 The Package shall be able to handle the integrated data objects corresponding to the observational programs carried out by ALMA. These objects may be implemented in any manner appropriate, though relations between the components of the object must be maintained through some mechanism. These include:

- **OL–3.1–R2.1** Program header information
- **OL–3.1–R2.2** Observation status information (and schedules themselves)
- **OL–3.1–R2.3** Field information
- **OL–3.1–R2.4** Data from interferometer and single-dish organized by:
  - **OL–3.1–R2.4.1** position
  - **OL–3.1–R2.4.2** subreflector state
  - **OL–3.1–R2.4.3** polarization products
  - **OL–3.1–R2.4.4** spectral channels
  - **OL–3.1–R2.4.5** frequency bands
  - **OL–3.1–R2.4.6** IFs
  - **OL–3.1–R2.4.7** subarray
- **OL–3.1–R2.5** Coherence function (visibility) data from interferometer, including:
  - **OL–3.1–R2.5.1** cross-correlations
  - **OL–3.1–R2.5.2** auto-correlations
  - **OL–3.1–R2.5.3** uncorrected and/or online WVR corrected (if chosen by user)
  - **OL–3.1–R2.5.4** phased-array data (if chosen by user)
- **OL–3.1–R2.6** Total power (autocorrelation) data from antennas in single-dish mode
- **OL–3.1–R2.7** Weights and/or data uncertainties
- **OL–3.1–R2.8** States indicating special modes (such as ON/OFF positions), for:
  - **OL–3.1–R2.8.1** OTF scanning
  - **OL–3.1–R2.8.2** subreflector switching
  - **OL–3.1–R2.8.3** frequency switching
- **OL–3.1–R2.9** Flagging data or masks
- **OL–3.1–R2.10** Diagnostic data and errors
- **OL–3.1–R2.11** A-Priori calibration data, including:
  - **OL–3.1–R2.11.1** bandpasses
  - **OL–3.1–R2.11.2** source flux densities
  - **OL–3.1–R2.11.3** antenna polarization leakages
- **OL–3.1–R2.12** Derived calibration data, including:
  - **OL–3.1–R2.12.1** gain tables
  - **OL–3.1–R2.12.2** flux bootstraps
- **OL–3.1–R2.13** Images and/or models produced from data
- **OL–3.1–R2.14** Processing history

Priority: 1

OL–3.1–R3 There must be a selection mechanism integrated within tools to choose between the various available data subsets such as:

- **OL–3.1–R3.1** sourcenames, and user specified data qualifiers (with wildcarding)
- **OL–3.1–R3.2** mosaic or scanning pointing centers

OL–3.1–R3.3 polarization products or channels
OL–3.1–R3.4 bands (frequency bands, IFs)
OL–3.1–R3.5 spectral channels
OL–3.1–R3.6 frequency-switched data
OL–3.1–R3.7 interferometer: subarrays
OL–3.1–R3.8 interferometer: WVR-corrected or uncorrected baselines

Priority: 1

OL–3.1–R4 Multiple pointing centers for mosaics must be supported.

Priority: 1

OL–3.1–R5 For polarization products, transformation must be provided to the desired Stokes output parameter(s).

Priority: 1

OL–3.1–R6 Averaging over time, bands and spectral channels shall be possible.

Priority: 1

OL–3.1–R7 Data taken in arbitrary (but parameterized) scanning patterns must be dealt with.

Priority: 1

OL–3.1–R8 The standard velocity frames must be supported:

OL–3.1–R8.1 heliocentric
OL–3.1–R8.2 geocentric
OL–3.1–R8.3 LSR

Priority: 1

OL–3.1–R9 Any existing flagging mask must be maintained and associated with the data it refers to during any subsequent operations (such as splitting of data sets). The mask shall be transferrable to other equivalent data sets (e.g. flags derived for a continuum dataset should be transferable to a line dataset derived from the same observations).

Priority: 1

OL–3.1–R10 Calibration and ancillary monitoring data (e.g. weather information, WVR data, pointing) that are provided in standard ALMA formats must be preserved, if requested by the user.

Priority: 1

OL–3.1–R11 Correlation products accumulated at multiple bit depths (16-bit,32-bit) or compressed data must be supported transparently (see ALMA SW-11 3.2.3-R5).

Priority: 1

OL–3.1–R12 Comprehensive and understandable processing history information for the data must be maintained and be exportable (both as tables and as plain text).

Priority: 1

OL–3.1–R13 Users shall have access to, and the ability to change, all aspects of the data including the header.

Priority: 1

OL–3.1–R14 The Package must support locking data files so that there is no possibility of one process corrupting a file that is also being written to by another process in the Package. The default model should be: "one writer, multiple readers."

Priority: 1

OL–3.1–R15 Merging (e.g. concatenation) and splitting of datasets shall supported.

OL–3.1–R15.1 Extraction of specified subsets of data (e.g. by source, time, subarray) shall be supported.
Priority: 1

OL–3.1–R15.2 Merging and reinsertion of data subsets (e.g. combination of different configurations, different epochs, mosaic pointings) shall be supported.
Priority: 2

OL–3.1–R15.3 The merging and splitting process, including selection of data to be merged (e.g. merging of frequency channels which may be differently labeled) shall be straightforward and not overly complex.
Priority: 2

OL–3.1–R15.4 The use of merged or split data in the Package shall be robust and transparent to the user. Subsequent operations shall work the same, whether the data is in its original form or built from merged data subsets, where possible.
Priority: 2

OL–3.1–R15.5 The appropriate calibration and ancillary monitoring data for the merged or split data (e.g. keeping only the data relevant to sources split out) shall be preserved.
Priority: 2

OL–3.1–R16 Distinctions between "single-source", "multi-source", single-dish, and interferometer datasets shall be avoided with context built into the dataset or header.

Priority: 2

OL–3.1–R17 Tools for access to header information, scan summaries (e.g. LISTR in AIPS), antenna locations (e.g. PRTAN in AIPS) shall be available, with printable output.

Priority: 2

OL–3.1–R18 When sorting or indexing is needed for performance enhancement, it shall be carried out in a manner transparent to the user.

Priority: 3

OL–3.1–R19 Compression of the data, with a selectable level of loss, shall be possible at various stages of the processing path.

Priority: 3

2.3.2 Data Import and Export

OL–3.2–R1 A variety of data formats must be supported by the Package.

OL–3.2–R1.1 The **ALMA standard archival data format** must be supported for both input and output without loss of functionality or information.
Priority: 1

OL–3.2–R1.2 Other standard formats **as designated by the ALMA Project** shall also be supported for both input and output without loss of functionality or information.
Priority: 1

OL–3.2–R1.3 Other popular data formats may be supported as the Package sees fit, possibly with some loss in information content. Flexibility and wider use beyond ALMA is to be encouraged.
Priority: 3

OL–3.2–R2 Disk and offline data storage (e.g. DAT, DDS, DLT) must be supported. **The project will maintain a list of media which the Package must support.**

Priority: 1

OL–3.2–R3 The Package must be able to handle, efficiently and gracefully, datasets larger than main memory of the host system.

Priority: 1

OL–3.2–R4 The ability to drop flagged data on export shall be included.

Priority: 2

OL–3.2–R5 The Package internal data format, which may be independent of other supported formats, must not be "bloated" and the required storage should not exceed by more than 1.5× the raw data format.

Priority: 2

2.3.3 Images and Other Data Products

OL–3.3–R1 Standard multi-dimensional images and arrays must be supported, including:

OL–3.3–R1.1 Spectra and image slices (1D)
Priority: 1

OL–3.3–R1.2 Planar images (2D)
Priority: 1

OL–3.3–R1.3 Spectral and Time Cubes (3D)
Priority: 1

OL–3.3–R1.4 Higher-dimensional Arrays (4D+)
Priority: 2

OL–3.3–R2 Other standard derived data products must be supported, including:

OL–3.3–R2.1 Point (CLEAN) models
Priority: 1

OL–3.3–R2.2 Pixel (gridded image) models
Priority: 1

OL–3.3–R2.3 Elliptical Gaussian models
Priority: 1

OL–3.3–R2.4 Uniform (optically thick) disk models
Priority: 2

OL–3.3–R2.5 Optically thin disk models
Priority: 2

OL–3.3–R2.6 wavelets
Priority: 3

OL–3.3–R2.7 Pixons
Priority: 3

OL–3.3–R3 Blanking of pixels shall be maintained through the processing of images.

Priority: 1

2.3.4 Foreign Data

OL–3.4–R1 Data produced by other interferometers and single dishes in similar observing modes shall be importable and processable if provided in **ALMA standard archival data format** or an **ALMA supported data format**

Priority: 2

OL–3.4–R2 Imaging data in standard formats (e.g. FITS) from astronomical instruments at different wavelengths shall be importable, with the ability to combine (coadd) these with ALMA data where appropriate. This should be through a set of widely used formats, with a **minimal list of supported standards established by the project**.

Priority: 2

2.3.5 Interaction with the Archive

OL–3.5–R1 Access from the archive (when such access is granted, e.g. when Package is run by ALMA staff) must be supported.

Priority: 1

OL–3.5–R2 The interface between the Package and archive must be able to provide data access (when such access is granted) without interfering with other access to the archive.

Priority: 1

OL–3.5–R3 Security and integrity of the archive must be ensured during these operations.

Priority: 1

2.4 Calibration and Editing

2.4.1 General Calibration and Editing Requirements

OL–4.1–R1 The Package must be able to reliably handle all **designated ALMA standard calibration modes** , possibly including but not exclusive to temperature controlled loads, semi-transparent vanes, apex calibration systems, WVR data, noise injection, fast-switching calibration transfer, planetary observations.

Priority: 1

OL–4.1–R2 Calibration, editing, flagging, and correction of data shall be easily reversible within the Package (ie. not requiring re-reading of the data from the archive).

Priority: 1

OL–4.1–R2.1 Logging of editing steps will be clearly marked in a history table or data object (possibly distinct from a more readable history).
Priority: 2

OL–4.1–R2.2 Individual edit undo is desirable.
Priority: 3

OL–4.1–R3 Data calibration operations shall take into account the scan structure and switching scheme of the data. The user shall be able to request calibration solution intervals that correspond to and reference from scan boundaries, for example.

Priority: 1

OL–4.1–R4 Data editing and flagging shall be possible based upon array, environmental, astronomical, and calibration monitoring data, including:

OL–4.1–R4.1 pointing data
Priority: 1

OL–4.1–R4.2 array tracking information (encoders)
Priority: 1

OL–4.1–R4.3 weather data (wind speed, temperature, relative humidity, pressure)
Priority: 1

OL–4.1–R4.4 T_{sys} data
Priority: 1

OL–4.1–R4.5 WVR data
Priority: 1

OL–4.1–R4.6 RFI monitoring
Priority: 2

OL–4.1–R4.7 site-test interferometer (STI) and/or tipping radiometer
Priority: 2

OL–4.1–R4.8 array monitoring points (e.g. dewar temperatures) if data provided in standard format
Priority: 3

OL–4.1–R4.9 other site instrument (e.g. FTS) providing data in standard format
Priority: 3

OL–4.1–R5 Data calibration, correction and flagging shall be possible based upon standard or user-defined models in either functional or tabular form. Arbitrary user-specified scaling of data shall be simple.

Priority: 1

OL–4.1–R6 Calibration shall involve flexible averaging of data and calibration quantities with user-controllable interpolation, filtering, weighting, and application scope.

Priority: 1

OL–4.1–R7 Interactive data editing, calibration, and display of calibration quantities shall be largely graphical and intuitive. Specialized editing display tools should include:

- **OL–4.1–R7.1** specification of data by selection on observtional parameters, and/or plotting versus these parameters, including:
  - **OL–4.1–R7.1.1** antenna
  - **OL–4.1–R7.1.2** baseline
  - **OL–4.1–R7.1.3** time range
  - **OL–4.1–R7.1.4** uv range
  - **OL–4.1–R7.1.5** pointing center
  - **OL–4.1–R7.1.6** azimuth, elevation
  - **OL–4.1–R7.1.7** hour angle range
  - **OL–4.1–R7.1.8** parallactic angle
  - **OL–4.1–R7.1.9** slice through data cube
- **OL–4.1–R7.2** displays of spectra and spectral cubes
- **OL–4.1–R7.3** display of closure quantities
- **OL–4.1–R7.4** display of and selection on monitor data quantities (e.g. Tatm, Tamb)
- **OL–4.1–R7.5** the above with baseline, time, band and/or channel averaging
- **OL–4.1–R7.6** for interferometer data, amplitude (phase) vs. time on each baseline (Difmap vplot), vs. time-baseline (AIPS TVFLG) with interactive zoom, selection, and clipping
- **OL–4.1–R7.7** editing of data points based on statistical quantities, including:
  - **OL–4.1–R7.7.1** a data point versus a running mean over a timescale
  - **OL–4.1–R7.7.2** a data point versus a median over a timescale
  - **OL–4.1–R7.7.3** an rms scatter in a time range
  - **OL–4.1–R7.7.4** difference versus a model
- **OL–4.1–R7.8** auto-scaling and user-specified scaling of axes
- **OL–4.1–R7.9** auto-scaled and user-specified colormap or greyscale
- **OL–4.1–R7.10** inclusion and marking of flagged data in plots and in auto-scaling

Priority: 1

OL–4.1–R8 Non-interactive and automated editing tools shall be available in the Package, including:

- **OL–4.1–R8.1** direct editing of data based user-specified ranges for quantities available for plotting or editing in interactive mode
  Priority: 1
- **OL–4.1–R8.2** automated editing with tunable criteria for automated selection of parameter ranges
  Priority: 2

OL–4.1–R9 Access to time history of calibration information (e.g. source catalogs containing flux density histories) shall be built into calibration engines. Output of calibration procedures shall be exportable into similar structures.

Priority: 1

OL–4.1–R10 The Package shall support the establishment and verification of the relative calibration of the various component epochs, configurations or other subsets for merged datasets.

Priority: 2

OL–4.1–R11 Data display and editing shall be effected through generic tools applicable to both single-dish and interferometer modes. These shall, as far as possible, present similar interfaces to the user and have the same look-and-feel.

Priority: 2

OL–4.1–R12 Editing shall be incorporated into most visualization tools where data or data-derived quantities are plotted, such as from calibration solutions, amplitude vs. uv-distance plots, or any number of other plots. A "see-it, flag-it" capability shall be the standard within the tools.

Priority: 2

2.4.2 Atmospheric Calibration

OL–4.2–R1 Atmospheric modelling shall be available in the Package. **The ALMA Project shall provide a standard model as used in the Pipeline processing of data which the package must support.** There shall be provision for the use of user-supplied models, or the package may have its own built-in models as an option.

Priority: 1

OL–4.2–R2 The Package shall be able to predict the absorption, emission and pathlength on the line of sight through the atmosphere at all ALMA bands using the model. The prediction will be based on the following data:

OL–4.2–R2.1 measured atmospheric parameters at the site: temperature, pressure, humidity
Priority: 1

OL–4.2–R2.2 measured atmospheric emission in the observed ALMA bands
Priority: 1

OL–4.2–R2.3 data from site test interferometer or tipping radiometer
Priority: 2

OL–4.2–R2.4 measured FTS data (if FTS available for ALMA)
Priority: 3

OL–4.2–R2.5 measured atmospheric profiles of temperature and water content if available from atmospheric sounders or other devices
Priority: 3

OL–4.2–R3 Atmospheric modelling shall be usable to derive by model fitting the system temperatures corrected for atmospheric absorption in all astronomical bands in use, in order to correct the observed amplitudes at various elevations.

Priority: 1

OL–4.2–R4 Atmospheric modelling shall be also usable to provide the conversion factors between WVR data and the water contribution to the astronomical phase in the astronomical bands.

Priority: 1

2.4.3 Interferometer Data

Note: Antenna-based determination of calibration quantities such as gains, polarization leakages, bandpasses, will be the primary form of calibration where appropriate. However, in addition to antenna-based calibration, baseline dependent corrections will be important in some cases. For example, coherence loss due to atmospheric phase fluctuation depends on baseline length (this aspect will be more important at higher frequencies) and must be taken into account if some of the WVR corrections are discarded while others are applied. Also, in general, the bandpasses are baseline dependent and contain non-closing terms.

OL–4.3–R1 Antenna-based determination of calibration quantities shall be available, and are the default choice for calibration in tools where appropriate, for quantities including:

OL–4.3–R1.1 antenna gains
OL–4.3–R1.2 polarization leakages

OL–4.3–R1.3 antenna-dependent bandpasses

Priority: 1

OL–4.3–R2 Baseline dependent corrections shall be supported for quantities including:

OL–4.3–R2.1 WVR corrections to subsets of baselines
OL–4.3–R2.2 closure errors
OL–4.3–R2.3 baseline-dependent bandpasses

Priority: 1

OL–4.3–R3 Gain corrections will be made based on differences between observed and modeled data quantities, possibly with iteration (e.g. self-calibration and determination of gains using calibration sources). Where solutions are discrepant or poor, automatic editing shall be possible.

Priority: 1

OL–4.3–R4 Calibration quantities (possibly stored in tables or data structures) shall be transferable between sources and/or frequency bands, after any necessary interpolation, extrapolation or smoothing. This will be the primary method of phase calibration transfer using fast-switching between source and calibrator.

Priority: 1

OL–4.3–R5 Determination of, correction for, examination of. and flagging based on closure errors shall be straightforward to carry out.

Priority: 1

OL–4.3–R6 Determination of the time-variable, complex bandpass using calibration source observations, and transfer to target sources, shall be simple and robust.

Priority: 1

OL–4.3–R7 Determination of polarization calibration quantities such as leakage (D-term or Jones matrix) and complex gain difference shall be an integral part of the Package, with the capability of performing full matrix calculations.

Priority: 1

OL–4.3–R8 Incorporation of standard models (e.g. planetary disks, models for HII region structure, known source spectra) shall be easy for calibration operations. User-supplied models in a standard format (e.g. ASCII table) for these quantities shall be supported.

Priority: 2

OL–4.3–R9 Redundancy (e.g. same, similar, or crossing baselines) shall be used wherever possible to increase accuracy of or to check calibration solutions. Editing based on this comparison shall be possible.

Priority: 2

OL–4.3–R10 Interferometric pointing, focus, baseline, and beam response fitting shall be available in the Package as a supplement to the on-line calibration.

Priority: 3

2.4.4 Single Dish Data

OL–4.4–R1 Straightforward and flexible fitting of spectral bandpass from calibration source observations is required.

Priority: 1

OL–4.4–R2 De-striping and adjustment of scan normalization factors must be available for single-dish OTF observations with overlapping and crossing scans.

Priority: 1

OL–4.4–R3 Calibration of system parameters such as temperature controlled loads and noise sources from observations of celestial sources shall be supported.

Priority: 3

OL–4.4–R4 Processing for pointing, focus, tipping, or beam-fitting data must be available for both single-scan calibration observations and for multiple datasets.

Priority: 3

2.4.5 Mosaicing Considerations

OL–4.5–R1 Determination of and correction for pointing offsets and the polarized primary beam is critical to the ability to reliably mosaic using ALMA, and thus must be available in the Package, preferably in several algorithmic forms.

Priority: 1

OL–4.5–R2 Careful cross calibration of the flux scales between ALMA interferometric data and single dish data is required for high fidelity imaging. There must be tools to cross-check and correct the relative calibration between mosaics and different component observations. *Note: this is particularly important and more difficult for ACA data.*

Priority: 1

OL–4.5–R3 The Package shall allow user-input of scan information given a parameterized pattern or interpolated from a provided list of pointing centers versus time

Priority: 3

2.4.6 Ancillary and Diagnostic Data

OL–4.6–R1 Output from the atmospheric monitoring (e.g. WVR, FTS) instrumentation provided in ALMA format shall be importable to the calibration software, for example to be used in flagging.

Priority: 1

OL–4.6–R2 Derivation of calibration and correction of data based on pointing, focus and subreflector status information shall be available in the Package.

Priority: 1

OL–4.6–R3 Environmental data such as weather (e.g. wind speed, temperature, dew point) supplied in ALMA format, as FITS extension tables, or as ASCII tables shall be importable for editing or calibration procedures, and easily incorporated into user-specified calibration models.

Priority: 2

OL–4.6–R4 Engineering monitoring information such as temperature sensor readings and tilt-meter outputs, included in ALMA format, as FITS extension tables, or as ASCII tables, shall be readable and incorporated into the calibration and editing process.

Priority: 3

2.5 Imaging

2.5.1 General Imaging Requirements

Because ALMA is inherently a multi-channel instrument, spectral cube mapping shall be built in as the primary mode from the beginning. Also, due to the high volume of data that can be produced by ALMA, it is imperative that the imaging and deconvolution tools in the Package be user-friendly, efficient, and flexible. This is the workhorse of the Package as far as most users will be concerned, and suitability and success of the Package will be judged with this in mind.

OL–5.1–R1 Imaging of data taken from any combination of ALMA exported data, the ALMA archive, or other instruments supporting common export formats must be provided. **A list of supported data and formats will be maintained by the project.**

Priority: 1

OL–5.1–R2 Efficient selection of subsets of the imaging data must be provided.

Priority: 1

OL–5.1–R3 Provision must be made for the utilization and development of a variety of imaging, deconvolution, and analysis algorithms, including:

OL–5.1–R3.1 raw ("dirty") images with selectable weighting (natural, uniform, Briggs robust)
Priority: 1

OL–5.1–R3.2 residual images after model subtraction
Priority: 1

OL–5.1–R3.3 single-scale CLEAN (Hogbom, Clark, Cotton-Schwab)
Priority: 1

OL–5.1–R3.4 maximum entropy method (MEM)
Priority: 1

OL–5.1–R3.5 linear mosaics
Priority: 1

OL–5.1–R3.6 non-linear mosaics
Priority: 1

OL–5.1–R3.7 non-negative least-squares (NNLS)
Priority: 2

OL–5.1–R3.8 multi-scale CLEAN
Priority: 2

OL–5.1–R3.9 modelfitting (point, Gaussian, disk)
Priority: 2

OL–5.1–R3.10 multi-frequency synthesis with different spectral models
Priority: 2

OL–5.1–R3.11 special function deconvolution (Pixon, wavelet)
Priority: 3

OL–5.1–R4 Image pixel and/or spectral channel blanking must be supported.

Priority: 1

OL–5.1–R5 Interactive graphical selection of deconvolution region masks shall be available.

Priority: 1

OL–5.1–R6 Multiple input datasets shall be supported directly in the tools, rather than requiring previous concatenation of the data.

Priority: 2

OL–5.1–R7 An integrated deconvolution, self-calibration, and editing/filtering tool shall be available, especially for novice users with data taken in commonly used modes.

Priority: 2

OL–5.1–R8 Images made on different equinox (e.g. B1950 vs. J2000) or different coordinate (RA,DEC and l,b) systems or different projection (tangent, sinusoidal, ...) shall be transformed, merged and compared appropriately.

Priority: 2

OL–5.1–R9 Image cubes using different velocity definition (optical or radio definition for Doppler velocity) shall be transformed and merged appropriately.

Priority: 2

2.5.2 Interferometer Imaging

OL–5.2–R1 High-fidelity imaging of the entire primary beam in all Stokes parameters is the primary goal — therefore, incorporation of the polarized primary beam response of the array is required.

Priority: 1

OL–5.2–R2 Imaging of direct polarization products (e.g. RR, LL, RL, LR) or Stokes polarization states (e.g. I, Q, U, V) must be selectable and interchangeable where possible given the data.

Priority: 1

OL–5.2–R3 There must be straightforward and seamless integration of data from multiple epochs and configurations.

Priority: 1

OL–5.2–R4 Simultaneous multiple-field imaging and deconvolution with gridded and ungridded (FFT and DFT) computation of model subtraction must be supported.

Priority: 1

OL–5.2–R5 Subtraction of continuum level from spectral data is required, in both the Fourier and image domain. In the case of uv-plane subtraction, flexible setting of the frequency channel ranges for the calculation of the continuum level and selection of the type of fitting (polynomial, spline) should be available in both GUI and CLI.

Priority: 1

OL–5.2–R6 There must be the ability to include "zero-spacing" values and short-spacing data taken in single-dish mode (both ALMA and non-ALMA data), with selectable weighting.

Priority: 1

OL–5.2–R7 There shall be the choice of FFT and DFT imaging (especially for small datasets).

Priority: 2

OL–5.2–R8 There shall be the provision for the near-field imaging of solar-system objects. This can be done through the introduction of phase corrections based upon the sphericity of the incoming wave front.

Priority: 2

2.5.3 Mosaicing and Single Dish Imaging Considerations

OL–5.3–R1 Combination of interferometer and single-dish data into mosaic imaging is essential.

Priority: 1

OL–5.3–R1.1 The ACA (if and when it is scoped as part of ALMA) must be integrally supported by the Package.
Priority: 1

OL–5.3–R2 Careful (polarized) primary beam correction and pointing correction is critical for high fidelity mosaic imaging and must be incorporated into the mosaicing algorithms.

Priority: 1

OL–5.3–R2.1 The primary beam calculation and correction must take into account the effect of on-the-fly scanning.
Priority: 1

OL–5.3–R2.2 A set of **ALMA standard beam images** will be made available by the project and distributed with the Package, with updates available for download when appropriate.
Priority: 1

OL–5.3–R2.3 The user shall be able to specify the primary beam in a number of forms, both analytic and tabular, in addition to the ALMA provided primary beam.
Priority: 2

OL–5.3–R3 Image plane and uv-plane gridding parameters and interpolation should be selectable based on desired image criteria (e.g. field-of-view, accuracy)

Priority: 1

OL–5.3–R4 Handling of interferometer (polarization product) and single-dish (polarization channel) data together may not be straightforward, and must be dealt with correctly by the Package.

Priority: 1

OL–5.3–R5 Scaling and de-striping of scans on image-plane combination shall be available.

Priority: 1

OL–5.3–R6 The Package must be able to produce an image by combining data observed on different rasters, possibly taken with different (regular or irregular) spacings and image centers.

Priority: 1

OL–5.3–R7 Pointing corrections (e.g. as determined by optical cameras or through monitoring data) shall be applicable to the data during imaging

Priority: 2

OL–5.3–R8 The imaging tools shall allow the option for the mitigation of the effects of non-coplanar baselines and sky curvature.

Priority: 2

2.6 Data Analysis

2.6.1 General Analysis Requirements

OL–6.1–R1 Seamless transformation between image-plane and uv-plane analysis is necessary.

- **OL–6.1–R1.1** Analysis based on goodness-of-fit to models, in both uv-plane (for interferometry) and image plane, shall be available.
  Priority: 1
- **OL–6.1–R1.2** Fourier transform of images between angular and uv domains shall be available.
  Priority: 2

OL–6.1–R2 Spectral and scanning baseline removal facility is required.

- **OL–6.1–R2.1** Polynomial baseline fitting shall be supported
  Priority: 1
- **OL–6.1–R2.2** Baseline fitting based on image statistics (e.g. median filtering) shall be supported
  Priority: 1
- **OL–6.1–R2.3** Fourier analysis of standing waves and their removal from spectra shall be available.
  Priority: 2

OL–6.1–R3 Translation between various astronomical quantities and units (e.g. Jy and K, MHz and km/s, R-J brightness temperature and true Planck temperature) shall be available.

Priority: 2

OL–6.1–R4 The astronomer must have the capability to develop their own tools or tasks, with easy access to data and images, and straightforward interface with the Package.

Priority: 2

2.6.2 Spectral Line Analysis

ALMA will be a spectroscopic machine, and thus the ease and flexibility of spectral line analysis is likely to be a driver for the Package. Automatic and user-controlled fitting routines will need to be included. Note that there is necessarily some overlap with the image cube analysis, but here we concentrate on issues relevant to traditional 1-D spectra.

OL–6.2–R1 Automatic and user-controlled measurement of line parameters shall be available.

Priority: 1

OL–6.2–R2 It shall be possible for the user to specify a velocity or frequency windowing region for line fitting.

Priority: 1

OL–6.2–R3 Available line fitting parameters and profiles shall include:

- **OL–6.2–R3.1** Gaussian line parameters (central and integrated intensity, line width, line center) for single or multiple lines
  Priority: 1
- **OL–6.2–R3.2** Damping profiles (Lorentizian)
  Priority: 2

OL–6.2–R4 A set of **ALMA standard line catalogs** shall be made available by the project and distributed with the Package, with updates available for download when appropriate.

Priority: 1

OL–6.2–R5 User importable line catalogs shall be supported by the Package as an ASCII table and in the **ALMA standard format**.

Priority: 2

OL–6.2–R6 Setting of fit constraints (e.g. spacing for multiple lines) shall be available and flexible (ie. by GUI or by parameter list).

Priority: 2

OL–6.2–R7 Export of fit results in ASCII-format is desirable.

Priority: 3

2.6.3 Image Cube Analysis and Manipulation

Because ALMA is inherently a multi-channel instrument, and due to the design of the ALMA correlator, spectral image cube can be considered to be the fundamental image structure. Single-channel or continuum images can be considered as subsets or instances of cubes. Note that the ability to pull lower-dimensional structures from larger-dimensional cubes is especially important.

OL–6.3–R1 The Package shall support the construction and analuysis of image cubes with a variety of axis choices, including:

OL–6.3–R1.1 time
OL–6.3–R1.2 position (RA,Dec and AZ,EL or general X,Y)
OL–6.3–R1.3 Fourier (u,v)
OL–6.3–R1.4 channel (frequency, velocity, band, IF)
OL–6.3–R1.5 polarization

Priority: 1

OL–6.3–R2 Basic cube rotation and transposition operations shall be available, including rotation not orthogonal to cube faces.

Priority: 1

OL–6.3–R3 It shall be possible to define multi-dimensional regions of the cube for further processing, both interactively and by specifying coordinates manually.

Priority: 1

OL–6.3–R4 Identification and reporting of image features (e.g. as determined in processing operations) shall be available, and interactive (where appropriate). Features shall include:

OL–6.3–R4.1 output in pixel coordinates (e.g. row, column)
Priority: 1
OL–6.3–R4.2 output in world coordinates (e.g. RA, Dec, Vel)
Priority: 1
OL–6.3–R4.3 convertible to other supported coordinate systems (e.g. precessed to B1950)
Priority: 2
OL–6.3–R4.4 exportable (e.g. as an ASCII file)
Priority: 3

OL–6.3–R5 The ability to extract same or lower-dimensional structures from higher-dimensional data cubes efficiently is required.

Priority: 1

OL–6.3–R5.1 Extraction of "cubical" sub-structures aligned with the original cube axes must be straightforward.
Priority: 1

OL–6.3–R5.2 User selection of extraction criteria must be possible through the GUI as well as scriptable.
Priority: 1

OL–6.3–R5.3 User-selectable sub-structures with arbitrary orientation within the parent cube, with appropriate transformation or interpolation, shall be possible.
Priority: 2

OL–6.3–R5.4 Extraction of data structures based on standard database (e.g. SQL) queries shall be available.
Priority: 3

OL–6.3–R6 The ability to collapse or integrate over sub-dimensions of data cubes in order to form "moments" is required. This shall be possible along any direction in the cube.

Priority: 1

OL–6.3–R7 The Package must have the capability of assembling lower-dimensional data structures into higher-dimension cubes.

Priority: 1

OL–6.3–R8 Blanking of pixels must be maintained through the analysis process.

Priority: 1

OL–6.3–R8.1 It must be possible to turn on and off different blanking (mask) levels, when blanking is set within the Package.
Priority: 1

OL–6.3–R8.2 Blanking shall not be destructive, and the original pixel value is retained (if defined).
Priority: 2

OL–6.3–R8.3 Interactive and automatic facilities for setting of any blanking parameters shall be provided.
Priority: 2

OL–6.3–R9 A variety of image processing and filtering operations on the cube shall be available, including:

OL–6.3–R9.1 Smoothing and convolution, with kernels including:

OL–6.3–R9.1.1 Uniform (box-car or top-hat) kernel
Priority: 1

OL–6.3–R9.1.2 Elliptical Gaussian kernel
Priority: 1

OL–6.3–R9.1.3 Symmetric polynomial
Priority: 2

OL–6.3–R9.1.4 User-supplied
Priority: 3

OL–6.3–R9.2 Clipping
Priority: 1

OL–6.3–R9.3 Windowing and boxing, with definitions of regions of interest recordable and passable for subsequent use
Priority: 1

OL–6.3–R9.4 Arithmetical operations, including:

OL–6.3–R9.4.1 addition of uniform level
Priority: 1

OL–6.3–R9.4.2 multiplicative scaling
Priority: 1

OL–6.3–R9.4.3 logarithm
Priority: 2

OL–6.3–R9.4.4 exponential
Priority: 2

OL–6.3–R9.5 Statistical operations over a defined region, including:

OL–6.3–R9.5.1 mean
Priority: 1

OL–6.3–R9.5.2 rms
Priority: 1

OL–6.3–R9.5.3 standard deviation from mean
Priority: 1

OL–6.3–R9.5.4 median
Priority: 1

OL–6.3–R9.5.5 mode
Priority: 2

OL–6.3–R9.6 Fourier and correlation operations on cube or cubical sub-regions, including:

OL–6.3–R9.6.1 Fourier transform
Priority: 1

OL–6.3–R9.6.2 power spectrum
Priority: 1

OL–6.3–R9.6.3 autocorrelation
Priority: 1

OL–6.3–R9.6.4 structure function
Priority: 2

OL–6.3–R9.7 Scalar arithmetic between different regions (including treatment of masked regions and differently shaped regions), including:

OL–6.3–R9.7.1 sum (difference)
Priority: 1

OL–6.3–R9.7.2 product (division)
Priority: 1

OL–6.3–R9.7.3 spectral index
Priority: 1

OL–6.3–R9.7.4 median
Priority: 2

OL–6.3–R9.7.5 maximum (minimum)
Priority: 2

OL–6.3–R9.8 Construction and comparison of vector quantities in the cube, including:

OL–6.3–R9.8.1 polarization (E) vector at each cube pixel
Priority: 1

OL–6.3–R9.8.2 rotation measure between different frequencies
Priority: 1

OL–6.3–R9.9 Vector arithmetic between different regions (including treatment of masked regions and differently shaped regions), including:

OL–6.3–R9.9.1 sum (difference)
Priority: 1

OL–6.3–R9.9.2 dot product
Priority: 2

OL–6.3–R9.9.3 cross product
Priority: 2

OL–6.3–R9.10 Interpolation across blanked or masked regions
Priority: 1

OL–6.3–R9.11 Fitting of models, shapes, profiles and functions over regions including:

OL–6.3–R9.11.1 polynomials
Priority: 1
OL–6.3–R9.11.2 exponentials
Priority: 1
OL–6.3–R9.11.3 Fourier modes
Priority: 1
OL–6.3–R9.11.4 elliptical Gaussian components
Priority: 1
OL–6.3–R9.11.5 uniform (multi-dimensional) spheres
Priority: 2
OL–6.3–R9.11.6 triginometric functions
Priority: 2
OL–6.3–R9.11.7 Lorentzian profiles
Priority: 2
OL–6.3–R9.11.8 user-supplied functions
Priority: 3

OL–6.3–R9.12 Filtering, including:
OL–6.3–R9.12.1 Fourier (Wiener) filtering
Priority: 2
OL–6.3–R9.12.2 Sobel filtering
Priority: 3
OL–6.3–R9.12.3 unsharp mask
Priority: 3
OL–6.3–R9.12.4 user-supplied filtering functions
Priority: 3

OL–6.3–R9.13 Deconvolution
Priority: 2

OL–6.3–R9.14 Resampling (e.g. at lower temporal or spectral resolution) after processing
Priority: 2

OL–6.3–R9.15 Calculus on scalar and vector fields in cube, including:
OL–6.3–R9.15.1 differentiation along paths
Priority: 3
OL–6.3–R9.15.2 integration over regions
Priority: 3
OL–6.3–R9.15.3 gradient of scalar fields
Priority: 3
OL–6.3–R9.15.4 divergence of vector fields
Priority: 3
OL–6.3–R9.15.5 curl of vector fields
Priority: 3
OL–6.3–R9.15.6 Laplacian of vector fields
Priority: 3

OL–6.3–R9.16 There shall be the capability to manipulate data cubes as general data structures, so that arithmetical and logical operations can be applied as object methods.
Priority: 3

2.7 Visualization

This is intended as the purely graphical part of data analysis. There is by necessity some overlap with the functionality discussed under Data Analysis, particularly that for image cube manipulation, and it would in fact be ideal if visualization and analysis were so closely integrated that there were no effective difference. The intention here is that the user is not only able to display pre-calculated images (processed using tools from the Data Analysis suite), but also has the capability of doing some processing and display on-the-fly as an integral part of the visualization.

2.7.1 General Visualization and Plotting Requirements

OL–7.1–R1 Plotting and display capabilities shall be integrated into the GUI tools throughout the Package.

Priority: 1

OL–7.1–R2 Standard type of plots must be supported, such as

OL–7.1–R2.1 X-Y plots
OL–7.1–R2.2 histograms
OL–7.1–R2.3 contour plots
OL–7.1–R2.4 vector plots
OL–7.1–R2.5 2D images
OL–7.1–R2.6 wireframe 3D surfaces

Priority: 1

OL–7.1–R3 "Blinking" between two different images, or stepping through a set of ordered images, at a user-selected rate and with adjustable transfer functions must be possible.

Priority: 1

OL–7.1–R4 Standard plotting formats shall be supported, both displayed (e.g. X window) and hardcopy.

OL–7.1–R4.1 There must be at least one designated standard output format (e.g. postscript) that can be converted by the user to a variety of formats using easily obtainable tools.
Priority: 1

OL–7.1–R4.2 The Package shall also support the output of a variety of commonly used formats such as FITS, postscript, pdf, gif and/or jpeg.
Priority: 3

OL–7.1–R5 Identification of cursor position shall be available for interactive plots. Where appropriate, this information shall be recordable and exportable. If you "see-it" you should be able to figure out where it came from.

Priority: 1

OL–7.1–R6 An extra "axis" of information shall be encodable on the standard plot types using color and/or intensity.

Priority: 2

OL–7.1–R7 The displays shall have similar look and feel to reduce the plotting learning curve.

Priority: 2

2.7.2 Display Appearance and Interactivity

OL–7.2–R1 Plot selection parameters (axes, limits, colormap) shall be conveniently controllable.

Priority: 1

OL–7.2–R2 There shall be interactive display zooming and unzooming capability within plot windows.

Priority: 1

OL–7.2–R3 The plot update speed shall not be a bottleneck. Speed shall be benchmarked, and should be commensurate with comparable plotting packages.

Priority: 1

OL–7.2–R4 Different line styles, sizes, thicknesses and colors must be available. User shall be able to manipulate intensity and color scales, interactively for graphical displays. The setup shall be saveable and reloadable.

Priority: 1

OL–7.2–R5 Basic axis transformations shall be built in to plotting, such as:

OL–7.2–R5.1 Logarithmic amplitude and intensity scale
Priority: 1

OL–7.2–R5.2 Different time and coordinate units and formats (e.g. hours, hhmmss, radians, ddmmss.s)
Priority: 1

OL–7.2–R6 User shall be able to augment plots and produce overlays of different data sets of standard formats.

OL–7.2–R6.1 Images with same axes, size and orientation shall be superposable directly, with basic control of colors and symbols.
Priority: 1

OL–7.2–R6.2 Overlay layer style shall be selectable, e.g. contours, greyscale, colormapped (RGB or HSV), or single color (i.e. one layer gets assigned intensity scales of red, another one of green, and one of blue).
Priority: 1

OL–7.2–R6.3 Overlay of axis grids shall be available.
Priority: 1

OL–7.2–R6.4 The user must be able to overlay functional fits (e.g. polynomials) or points read in from standard tabular files
Priority: 1

OL–7.2–R6.5 It shall be possible to place data sets in "layers" be which can be interactively colormapped, and switched on and off.
Priority: 2

OL–7.2–R6.6 It shall be possible to display and overlay data with different coordinate systems, i.e. the coordinate system of the display can be chosen independent of the system the data were observed in and the data transformed appropriately with pre-computation.
Priority: 3

OL–7.2–R6.7 It shall be possible to shift, rotate and scale the images interactively.
Priority: 3

OL–7.2–R7 Users shall be able to synchronize multiple display windows, for example such that zooming to a given pixel in one image window will select the equivalent pixel in the slaved windows.

Priority: 2

OL–7.2–R8 Users shall be able to add annotation, both interactively and through scripts, including text with various fonts (including Greek letters), symbols (e.g. all the symbols provided by the LaTeX package with AMSTeX extension), arrows, geometrical figures like boxes and circles, etc.

Priority: 2

2.7.3 Visibility Data

OL–7.3–R1 Plotting of commonly used basic data and calibration quantities must be straightforward and easily accessible from all relevant tools. Any quantity should be displayable versus any other (or any two plus another encoded as a z-axis or intensity) as long as these quantities are defined for the same visibility or calibration solution interval.

These displayable quantities include:

OL–7.3–R1.1 real and/or imaginary
OL–7.3–R1.2 amplitude and/or phase
OL–7.3–R1.3 delay and/or rate
OL–7.3–R1.4 time
OL–7.3–R1.5 hour angle
OL–7.3–R1.6 parallactic angle
OL–7.3–R1.7 uv distance
OL–7.3–R1.8 u and/or v
OL–7.3–R1.9 baseline
OL–7.3–R1.10 channel or frequency or velocity
OL–7.3–R1.11 azimuth and/or elevation

Priority: 1

OL–7.3–R2 Data selection parameters shall be understandable (e.g. by antenna name or number instead of antenna table entry number, polarization name RR or I) and straightforward, using graphical browsers (in GUI mode) and/or standard selection language (e.g. SQL queries) in script mode.

Priority: 1

OL–7.3–R3 It shall be possible to interpolate or extrapolate any tabulated quantity onto a visibility or calibration solution point, and then manipulate these like extra visibility information.

Priority: 3

2.7.4 Other ALMA Data

Although the Package will not likely be the primary vehicle for the ALMA staff to assess the state of the array, it is intended that users (as well as staff) have the full capability of using ancillary data provided by ALMA to aid in the processing and understanding of their data. Therefore, the Package should be able to deal with this data in as user-friendly a manner as possible.

OL–7.4–R1 The Package shall be able to plot standard ALMA-format ancillary data, including

OL–7.4–R1.1 amplitude or single-dish power versus AZ and EL
Priority: 1
OL–7.4–R1.2 focus data and curves
Priority: 2
OL–7.4–R1.3 pointing data and offset vectors
Priority: 2
OL–7.4–R1.4 WVR output data
Priority: 3
OL–7.4–R1.5 holography and beam map data
Priority: 3
OL–7.4–R1.6 monitor point values (e.g. temperatures)
Priority: 3

2.7.5 Image-cube Manipulation

OL–7.5–R1 Overlay of selectable coordinate grids (e.g. J2000, B1950, galactic, ecliptic, pixel number) must be available. It must also be possible to overlay multiple grids.

Priority: 1

OL–7.5–R2 Histograms of pixel values must be easily produced for selected regions of the cube.

Priority: 1

OL–7.5–R3 It shall be possible to view subsets or slices of data cubes

OL–7.5–R3.1 for axes aligned with cube faces
Priority: 1
OL–7.5–R3.2 for arbitrarily aligned axes
Priority: 3

These shall be interactively selected in GUI mode.

OL–7.5–R4 It must be possible to plot values of the same pixel in different cube layers, or in different images, against each other.

Priority: 1

OL–7.5–R5 Data cubes must be viewable as movies with selectable frame rates and layer step directions.

Priority: 1

OL–7.5–R6 Interactive display of spectra corresponding to a pixel or region in a displayed image shall be supported.

Priority: 2

OL–7.5–R7 Interactive display of a 1D slice taken from a 2D image, such as dragging the line on the map to bring up a position-velocity diagram, shall be available.

Priority: 2

OL–7.5–R8 Plotting of spectra on a pseudo-grid corresponding to position on a raster (e.g. a "stamp map" or "profile map", basically thumbnail spectra in panels corresponding to position) shall be possible.

Priority: 2

2.8 Special Features

2.8.1 Simulation

Note: Inclusion of at least moderate simulation capability in the Package will benefit the user who may not have access to the more comprehensive simulator built into the Pipeline and Online systems. In the spirit of providing the user with offline options equivalent to those available in the project, it is desirable that some simulation capability also makes it into the Package albeit at low priority. The main goals would be to allow the user to simulate basic ALMA modes based on input models or images and to replace existing data in ALMA format with simulated data.

OL–8.1–R1 The Package shall inherit the basic simulation capability of the ALMA Online Data Processing Pipeline (Levels 1 and 2) as outlined in ALMA Memo SW-11, as per (5).

Priority: 2

2.8.2 VLBI

Note: It is assumed that VLBI data will be reduced in whatever package supports general VLBI observations, as ALMA will not be a stand-alone VLBI instrument. There are no VLBI requirements on the ALMA Offline Data Processing Package, though it may end up being the case that the Package also supports VLBI processing. This would have some advantage to users of the ALMA Package, and thus it is worth including a low-priority requirement to this effect.

OL–8.2–R1 It is desirable that the Package also support general VLBI processing suitable for dealing with ALMA data, in order to allow general ALMA users access to VLBI science without having to learn a completely new software package.

Priority: 3

2.8.3 Solar System Object Observing

The Sun and planets are likely to be important and intersting targets for ALMA observing. The requirements are likely to be strongest on the actual hardware (e.g. high frequency and time resolution) but software compatibility must also be considered. In particular, solar and planetary observations require a special effort in tracking and handling of ephemerides.

OL–8.3–R1 The Package must be able to calculate and compensate for the position of moving objects in the solar system.

OL–8.3–R1.1 Ephemerides must be provided for major solar system objects, including:

OL–8.3–R1.1.1 Sun
OL–8.3–R1.1.2 Moon
OL–8.3–R1.1.3 planets
OL–8.3–R1.1.4 major asteroids (all known with 50km dia. greater)
OL–8.3–R1.1.5 known short period comets

Priority: 1

OL–8.3–R1.2 Ephemerides must be calculated in available reference frames and equinoxes, including:

OL–8.3–R1.2.1 topocentric coordinates
OL–8.3–R1.2.2 geocentric coordinates
OL–8.3–R1.2.3 J2000 (and other FK5 equinoxes)

OL–8.3–R1.2.4 epoch of date
Priority: 1

OL–8.3–R1.3 The Package shall be able to import a user-supplied ephemeris in tabular form.
Priority: 2

OL–8.3–R1.4 The Package shall calculate positions from user-provided orbital elements.
Priority: 2

OL–8.3–R2 The Package must carry out astrometry for moving sources.

Priority: 1

OL–8.3–R3 The distance visibility amplitude correction for distance (from ephemeris for major objects, or user-supplied) shall be computed.

Priority: 1

OL–8.3–R4 The Package must calculate quantities for the "physical ephemeris" for the enumerated major solar system objects, including:

OL–8.3–R4.1 subearth latitude and longitude
OL–8.3–R4.2 subsolar latitude and longitude
OL–8.3–R4.3 position angle of North Pole
OL–8.3–R4.4 season
OL–8.3–R4.5 phase angle

Priority: 2

OL–8.3–R5 The Package shall do coordinate transformations from plane-of-sky to planetocentric for a selection of projections, including:

OL–8.3–R5.1 cylindrical
OL–8.3–R5.2 mercator
OL–8.3–R5.3 stereographic

Priority: 2

OL–8.3–R6 The imaging corrections for the near-field shall be included.

Priority: 2

OL–8.3–R7 In addition to the standard models available in the data analysis tools, there shall be provision for:

OL–8.3–R7.1 prolate and oblate ellipsoids
OL–8.3–R7.2 limb-darkened disks (polynomial, $\cos^n$, Legendre polynomial)

Priority: 2

2.8.4 Pulsar Observing

ALMA's pulsar needs will be much easier to satisfy than those of centimeter-wave arrays since dispersion smearing is a non-issue at these frequencies. We expect the main out-of-the-ordinary requirement will be integration into multiple phase bins. A minimum of 16 phase bins would be scientifically interesting. Pulsars have periods as short as 1.6 ms, so a minimum bin size of 0.1 ms would probably suffice. The main use of this binning capability will likely be measuring the on-pulse and off-pulse brightness which can be used to estimate or place bounds on the temperature of the neutron star. Pulsar timing is unlikely to be done with ALMA, barring the installation of specialized hardware. It is conceivable that ALMA would be a great instrument for pulsar parallaxes, especially if pulsars are detectable at 900GHz, and will likely be the best instrument in the southern hemisphere for years to come. The minimal spec of 16 pulsar phase bins would be sufficient for this work.

OL–8.4–R1 The Package shall support the processing of time-series data that has been folded into multiple phase bins. At least 16 such bins shall be supported.

Priority: 3