**eDisk Data Reduction Workflow**

**Data Collection and Preparation (Senior Reducer)**

1. Upon delivery of data from JAO, senior member of data reduction team (Tobin, Sheehan, others to be named) download data, restore data to calibrated state by running scriptForPI.
2. Weblog from calibration pipeline is inspected
   1. Check hifa\_gfluxscale - ensure no systematic offsets between spws
   2. Check hifa\_applycal, spws 33, 35 ,37 specifically to ensure that phase and amplitude versus time and frequency of calibrated data are flat
   3. Check images for phase calibrator and check source (if present) to verify consistency of peak intensity in each spw.
   4. Check ALMA calibrator catalog to see how well-monitored flux calibration source is.
3. Inspect ALMA standard data products to see if source is line rich/line poor and whether there is lots of extended emission or mostly compact
   1. Advise reducer of these features
4. Make tar.gz of the ALMA PL calibrated MS file(s) for full calibrated dataset
   1. The directory should have the following structure:

SourceName/

|--- SB/

|-------- SB\_dataset\_1.ms

|-------- SB\_dataset\_2.ms

|-------- …

|--- LB/

|-------- LB\_dataset\_1.ms

|-------- LB\_dataset\_2.ms

|-------- …

1. Prepare ERDA data repository directory structure if it does not exist
   1. Create directory structure if it does not exist

edisk/data/SourceName/

|--- ShortBaseline/

|------------------------ALMA\_pipeline\_calibrated\_data/

|---------------------------------------------------------------- SB\_dataset\_1.ms

|-----------------------------------------------------------------SB\_dataset\_2.ms

|------------------------Standard\_ALMA\_Products/

|------------------------eDisk\_Image\_Products/

|------------------------eDisk\_calibrated\_data/

|----------------------------------------------------SB\_dataset\_1\_cont.ms

|-----------------------------------------------------SB\_dataset\_2\_cont.ms

|-----------------------------------------------------SB\_dataset\_1\_line\_contsub.ms

|-----------------------------------------------------SB\_dataset\_2\_line\_contsub.ms

|------------------------scripts/

|--- LongBaseline/

|--- <Repeat short baseline>

* 1. Upload tar.gz of ALMA PL calibrated MS files(s) to ERDA data repository
  2. Upload ALMA standard image products to ERDA data repository

1. Data reducer assigned

**Self-calibration and Imaging (Junior Reducer)**

1. Download standard processing scripts from <https://github.com/jjtobin/edisk>
   1. git clone <https://github.com/jjtobin/edisk>
   2. If you already have cloned, run ‘git pull’ to update
2. Download eDisk CASA 6.2.0 package **ONLY USE THIS CASA!!**
   1. <https://astrocloud.nrao.edu/s/y9eQodm7ZMrepZf>
   2. ERDA location: eDisk/software/casa-6.2.0-124-eDisk.tar.gz
   3. Includes analysisUtils scripts and astropy in tarball
   4. Open casa and run ‘!update-data’ to update CASA data repository
3. Download ALMA PL calibrated MS file(s) and untar.
4. Use the two scripts labelled generic\_\* as a templates; copy them into the reduction directory.
5. Start with \*\_continuum\*.py reduction script and run interactively by copying and pasting commands into CASA terminal.
   1. Edit relevant portions of data reduction script . There are ONLY four places where the script should be edited:
      1. STARTUP/METADATA SPECIFICATION:
         1. **WD\_path:** Set to the directory that came from the tar file with the data.
         2. **pl\_data\_params:** Dictionary containing the locations of the raw datasets.
            1. May need to be updated with additional entries for each execution.
         3. **data\_params:** Dictionary containing relevant information for each track.
            1. Data\_params dictionary may need multiple blocks inserted for each dataset depending on number of executions.
         4. **Initial Imaging For alignment:** After images are made for each execution and Gaussians are fit. Pick best image to be reference image for alignment. If long baseline data are included, use a position from a long baseline dataset. **Manually fill-in a value for data\_params[i]['common\_dir'] from the best image.**
      2. SELF-CALIBRATION PREPARATION:
         1. **mask\_pa, mask\_maj, mask\_min:** Set such that they specify an ellipse that contains the bulk of the emission from the source in the image. Used to get SNR statistics
      3. SELF-CALIBRATION:
         1. Adjustments to specific self-calibration clean depths (nsigma) and solution intervals (solint) may be needed
         2. Adjustments to the image size may be needed, smaller is ok if sources are at center of field and no additional source present
         3. Adjustments to number of iterations of self-calibration to be done (e.g. some sources cannot do per-integration self-calibration, stop earlier. See IRAS 15398 as an example)
            1. Amplitude self-calibration can be ineffective/detrimental for certain datasets (IRAS 15398 as an example)
      4. FINAL IMAGING
         1. Adjustment of the image sizes may be needed here for both the final continuum images and the final spectral line images
   2. Be sure to check results as you go, in particular:
      1. Checks on Gaussian fitting a recentering of data
      2. Checks on the rescaling of data to have equivalent flux density scale
      3. Data and residual images after each self-calibration step
      4. Amplitude and phase self-calibration solutions after each self-calibration step
6. After interactive run of script, re-run non-interactively in a fresh directory to reproduce interactive results
   1. Script should run to completion non-interactively
7. Edit and run the \*\_spectral\_line\_\*.py reduction script on the data.
   1. There is ONLY one place where edits should be made to this script:
      1. STARTUP/METADATA Specification:
         1. **WD\_path:** Set to the directory that came from the tar file with the data.
         2. **Automasking parameters:** These can be adjusted depending on how well automasking does in the particular images. Two examples that work well for compact and extended emission are provided.
8. Copy plot\_final\_images.py script to the working directory to make a single pdf containing images of all of the data products generated.
9. Script is then sent to senior member of the reduction team via e-mail/Slack (Tobin, Sheehan, etc.) who will check the script and re-run the script non-interactively to verify results. And data cubes from reducer’s run will be uploaded to temporary storage on ERDA site.
   * 1. Script may go back to reducer for more work if necessary
     2. If accepted as-is or only minor changes, script will be committed to github repository and senior team member will upload the products to the ERDA repository
10. Data reduction marked as complete in spreadsheet by Senior member