

MAST Users Group Report – July 2009

The MAST users group (MUG) met at the Space Telescope Science Institute on 9 July 2009. MUG members are Mike Crenshaw (Georgia State University), Steve Howell (Chair, NOAO), Duilla deMello (CUA), Casey Papovich (Texas A&M), Evgenya Shkolnik (DTM, replacing C. Chen), and Ben Williams (University of Wisconsin) and all were present at the meeting. This was the first MUG meeting for Crenshaw, Shkolnik, and Williams and the last for deMello. One additional MUG member will need to be recruited for next year.

Overall Comments

The committee notes that the presentations were good overall: well prepared and informative. The MUG was very impressed the MAST archive, in particular the scientific productivity of the archive and its excellent NASA senior review.

Papers from archival data are now dominating the HST publications providing a strong metric of the value of the archive. The Hubble Legacy Archive (HLA) is proceeding well and the MUG was happy to see the wide array of scientific products that result from this effort. The MUG was happy to see the transition from the Sybase to the SQL servers. The fact that the improved SQL system will be 10x faster in CPU time and use 6x less memory is a big plus.

The initial work on collating and archiving JWST initial products such as calibration data is a worthwhile activity. General new mission work is proceeding well and the MAST staff has a good grasp on the balance between older mission data and new starts.

The MUG recognizes the usefulness of the MAST survey to the MAST team. A suggestion for the survey part related to MAST archive work may be to ask for rankings of desired additional features in addition to whether they find them useful or not, not just "would you like to see XYZ implemented. The MUG suggested a MAST forum similar to that of the IRAF group (See <http://iraf.net>) as a way to record and publicize common questions to the MAST Helpdesk. We were surprised to find that MAST has already done this and a link to it appears right on the MAST pages. Perhaps a larger link or more advertisement of it somehow may be in order. It would be helpful for someone to comb through the messages sent to archive@stscu.edu and other "help" messages/calls to identify important and/or common concerns and add these to the forum. In addition, future questions sent directly to the MAST Helpdesk could be answered within the Forum with its link sent to the questioner.

MAST might consider some document or presentation materials be developed to highlight MAST as the example for other national astronomy archives, including ground-based facilities. This is beyond the scope of MAST, but maybe there is some way for MAST to influence the funding of improved archives for the ground-based facilities (specifically NOAO and NRAO).

Hubble Legacy Archive

The MUG believes that the HLA is a great endeavor for MAST to undertake. We suggested last year to advertise it further and the MAST has been responsive to this task. A further suggestion this time is to disseminate the fact that many science products are available in the HLA not just “pretty pictures”. MAST needs to highlight the science products in example form; produce example “cookbooks” for a few specific projects that can be accomplished with HLA products. The MAST should look into holding a HLA 1-2 day “school” during which participants are shown the wide array of products and led through example projects. The MUG believes in the educational use of HLA products as well as mainstream science, and suggests that MAST work with the Institute outreach group in this area.

The documentation of the HLA details needs to catch up with the archive products so that scientific use can be increased. Specific to HLA, the documentation should extend to all catalogs, DAOPHOT, SExtractor, etc. The exact parameters (configuration files, etc) should be stated clearly and obtainable by the user. Any weight (exposure-time) maps, or “rms” maps should also be available. It may have been useful for the MUG to see a live demo of the new data products and search tools available in the HLA. Items such as “DADS” format need to be explained well in linked documents. We commend the MAST for their work on the HLA and look forward to the publication of a PASP paper detailing the archive products within the next year.

Some specific notes about the HLA interface and items to consider are:

- 1) Footprint polygon downloads (as ds9 regions files perhaps?)
- 2) Add ability to refine search with more than one filter chosen in the spectral column. Right now, it’s either all filters or a single filter.
- 3) Get original data for footprints (flt and/or crj images) from HLA for specific fields defined in the footprints and inventory section. Perhaps this could be a different data “level” (level 0?)
- 4) Link parallels to primaries so the Target names can be meaningful for data from parallels.
- 5) Mark search coordinates in HLA footprint view
- 6) The production of nice color mosaics; although these may not be as valuable scientifically as more flexible search and retrieval techniques. We realize this is an ongoing effort at the present time.

GALEX

GALEXview is a very nice tool but the MUG impression was that its default mode was a bit too complex for the first time user. The default mode could be set to a very simple interface, for example including an angular scale with none of the boxes checked. A simple cookbook for GALEXview would be very useful. Items such as “show all sources” (e.g., not returning clearly bad sources or non-detections, bad edge sources, etc.) and similar could be options but not the starting default. The user can then add complexity as they become familiar with the interface and GALEXview tools available. The MAST has a “to do” item of personalized login pages allowing users to save their interface mode, so for GALEXview this would be saved at the complexity level of the user.

There was some discussion as to who should be developing tools such as GALEXview; MAST or VAO? Perhaps MAST should explore cost-sharing for such endeavors. Overall, the MUG believes that the style, performance, and interface of GALEXview are a good match and provide general methods likely to be useful for other MAST products.

Some specific GALEXview suggestions are:

- 1) need to deal with the offset by the NUV and FUV hits of the AIS. Appears as two different objects but are really from the same object. One way to correct for this is that one needs to rerun a FUV search for those with NUV hits with some small search radius. However, there is no way to then identify FUV-only hits.
- 2) Output of CasJobs for GALEX queries should be the same as that returned by GALEXview. Otherwise, it hard to port data files.
- 3) a column with a quality flag is essential, with just those that are certainly spurious detections just removed completely from the returned table.
- 4) a classification flag is valuable as well (star/galaxy/other)
- 5) include the GALEX nearby galaxy catalog (integrated photometry of ~1000 galaxies) into GALEXview.

VO Activities

The MUG was happy with the VO interfacing and archival movement to the VO system. New missions should be required to deliver their science products to MAST in a VO compatible format agreed upon with MAST approval. MAST should not be required to put effort into converting these data in to VO format. The spectral containers project is highlighted here as one example of the fine work MAST is doing. They are leading the way in some specific VO format definitions and are clearly a world leader in the VO project. The MUG was assured that the MAST was aware of being careful to not duplicate efforts between MAST and VO activities.

Kepler Mission

The MAST has made great progress in preparation for the data that will be received from this mission. Kepler is one example of the difficulties MAST has at times in receiving full and proper documentation in a timely manner from a mission. The needed documentation tends to arrive eventually and the MUG would suggest that NASA sets more rigid requirements on future missions in this area as the MAST team should not have to be the enforcers of archive policy related to documentation nor have to spend time bugging the mission to deliver such work.

The Kepler archival data product of most interest to MAST users will likely be the time series light curve product. The MAST needs to work with the mission and the Kepler Data Management Center (DMC) at StSci to establish proper and common values for these data products in the area of time and “flux” units. Header values in all products, light curves and image data, should be in agreement and equal across all archive products. We point out here that these issues are mainly mission related and not MAST deficiencies. Light curve data should be available to MAST users as FITS tables and ASCII files, both easily downloadable. A quick look, simple plot tool (preview) would be useful for archive light curves to allow the user to examine potential data sets before download. A tool to search the light curves with user specified parameters (such as all light curves with variability greater than 0.5 magnitude, or all light curves with g-r color less than 0.4) would be useful as well.

MAST Documentation

The MUG realizes the large and complex issue that user documentation represents. We applaud the MAST for its good work in that area. Our suggestions to the MAST for documentation are varied in scope, some small and others large, and we fully expect that they all will not be possible to complete.

Outdated documents are often useful to researchers however, these need to be clearly identified as old and attempts should be made “hide” them upon initial searches -- Possibly have a MAST “documentation” web page with links to current and older versions of documents, organized by mission or dataset. All documents should be available in .pdf and HTML formats. The MUG viewed video tutorials as a low priority.

Future MAST plans

The MAST presentations included a number of future plans for the archive. The HLA footprint project, especially in connection with other missions, was viewed as a high priority product. Adding additional space missions to the footprint project will be valuable as well. The documentation of details of the HLA science products was also thought to be a high priority especially the input parameters and methods of archive product production. The ability to search the MAST archive products for solar system

(i.e., moving objects) was deemed a good extension for MAST and likely to be a highly desirable science product.

If MAST implements a personalized "login page" for all users it could include previous queries of MAST, and previously selected default parameters. Users can also sign up for emails regarding updates to data reduction, etc. If a registered-user system is implemented:

- 1) a user could be alerted when an updated reduction of the archived data he/she has previously downloaded is released. (if they check off some box during registration.)
- 2) a user could receive a keyword or target alerts which would be very helpful. i.e. if new data is uploaded to MAST (or other archives) about my favorite object, I'd like to receive an email about that with a link to the data (or the query form).
- 3) personal preferences could be stored.

The "Keyword" search tool project was slightly unclear to the MUG but seems to hold promise as a highly useful tool. We look forward to a progress report next year. The MUG notes that the original HST keywords, which changed some from cycle to cycle, should be kept. But there should be a new keyword based on a uniform system that collates all of these. The IUE and FUSE categories were actually not too bad. This keyword search project will be of relatively high priority, for those people that want to search for all objects of a given type in the archives. The literature-based approach sounds interesting, and should be investigated more.

In these days of large on-line data sets and big surveys with archive data available now and more to come, the MUG would like the MAST to explore a method by which the maximum submitted targets can be made greater than 1000 (up to ~100,000). Perhaps a user has to demonstrate their search is valid using a smaller number of targets after which the maximum submitted target limit is lifted temporarily for some time period. We leave the details of this idea to the experts at MAST.