

MAST Users Group Report – September 2011

The MAST users group (MUG) met at the Space Telescope Science Institute on 23 Sept. 2011. MUG members are Mike Crenshaw (Georgia State University), Steve Howell (Chair, NOAO), Casey Papovich (Texas A&M), Megan Donahue (Michigan State University), Ben Williams (University of Washington), and Jay Andersen (STSci). Donahue attended by telecom/Webex. This was the first MUG meeting for Donahue, and all remaining members, sans Andersen, have reached their term limits.

Overall Comments

We would first like to commend the MAST on their 1st place standing in the 2011 Senior Review. A testimony to their continued ability to make themselves highly valued by the community. Great job! We appreciate the effort and level of presentation delivered to the MUG during our review. We found them uniformly excellent and inspiring.

The MUG continues to be impressed with the product output of the MAST group. The MAST runs at a high level of products and reliability especially given its small budget and very busy staff. Current holdings are near 200TB and projected to be near 500 TB by 2015. Kepler is the current fastest growing data set. The MAST archive was again shown to be of great benefit to the astronomical community and is clearly a very important tool as evidenced by the large usage in terms of registered users, data delivered, and papers produced. About 700 papers annually contain archive data, a number that exceeds PI-type papers for all served missions. It would be interesting to have some metric to know how the high-level-science data figures into the increasing archive-related publication.

MAST staffing at a flat level for the next few years, is a concern to the MUG. It may severely limit value added tool development for missions such as Kepler and ACS.

MAST has been very responsive to the MUG and user comments, concerns, and suggestions for new or better interfaces. The MAST forms and documentation pages are a prime example of this. The “single line” interface is grand. The MUG realizes that documentation is a large, moving target with version control, and delivery of documentation from outside missions being areas of issue. When missions end, MAST tries hard to make sure the mission knowledge can be captured. Need to make a documentation interface as there is still no easy place to find all latest and greatest documentation.

The move of the archive and data service to Linux is impressive. This has made the entire archive path from form input to data served much faster (about 6.5 times) and was done with no user hiccups or downtime.

The loss of the ST-ECF center is non-negligible. Some expertise, archive help, and new software projects will be lost. Spectral expertise (esp. grism) is a specific area. Both MAST and ST-ECF did a good job to make sure that final software projects were finished and much of the information and archive products were transferred to MAST before ST-ECF went off-line.

The MUG notes that the MAST are anticipating more high-level products from the community, and that they are finding ways to be pro-active about how to accept such data so that the community can benefit. It may be too much at this point to require all data products delivered to be in a formal VO format, but MAST is in the best position to see that the data satisfy some minimal standards of uniformity.

We once again suggest that NASA HQ place tighter guidelines on missions for their delivered products in accordance with common and best practice archive data. MAST and the user community would benefit greatly from mission data delivered in the proper format the first time.

The MUG suggested that large proposals for missions such as HST should be highly encouraged (or required) to produce high-level archive products from their data. The "requirement" could be levied during the funding negotiation and be formal. MAST will clearly be ready to provide input about how the deliverables in such "contracts" or other agreements should be specified.

Overall, the MUG was impressed to see the collaboration that goes on between MAST, the HLA, and the VO. Rather than competing, they truly are collaborating, and they make an effort to come together whenever possible, creating common interfaces that are more useful than either interface alone.

HLA

The Hubble legacy archive continues to be a top-notch archive product. The interface and plotting tools are really nice to use. The new data formats allow for better use as well and the new WFC3 pipeline being near completion is noted. The development of the footprint service now being adopted by Chandra shows the great utility of this product and attests to its ease of transport to another archive site and the forethought in its development.

The HLA's image-viewing interface has been evolving very nicely. It helps astronomers enormously to preview many different aspects of a dataset without having to first download them in fits format.

The MUG hopes that the MAST archive spectral data can be brought into the same high level of content which the HLA has done for image data. We realize this is a big job and we encourage the MAST to collaborate with experts such as those at GSFC and U. Colorado.

The all-sky "source lists" is another large project (few years) but again the MUG feels this will be a good addition to the HLA toolbox, In order to have the available resources to perform this work, we suggest that the MAST prioritize its "future" list presented at the MUG. Some items may fall off the table, but that could be the price to finish the higher priority tasks.

The MUG suggests an easy link from the inventory list (perhaps simply an extra column) to the details to allow the individual input fit-level exposures of each drizzled image to be readily available to the user.

GALEX

The GALEX mission is ended in a rather quick fashion. Mission support for the instruments, data delivered and general knowledge of the spacecraft will soon be hard to obtain or lost forever. MAST is aware of this and working to gather documentation and final data products while they can.

The general development of GALEX archive tools is going well and most users are quite happy with the services offered. The MUG believes that it would be nice if GALEXView reported an upper limit (perhaps a simple flux per basic point source aperture) when no sources are returned within the user search radius IF the coordinates fell within the GALEX coverage. Lower down on the wish list, we suggest that a flag be returned that notes when a source or a search aperture is close to an edge.

The GALEX photon lists project seems to be a good idea to the MUG so that time-resolved photometry could be produced. Getting this data set to archive, while available, makes sense. However, this archive product appears only to be of use if a tool to produce such light curves is developed and made available to MAST users. The photon lists should also be linked to the GALEX point source catalog. This work would likely make a good ADP proposal.

SWIFT

Adding the SWIFT UVOPT images to the MAST archive, particularly in order to allow them to be a part of the general footprint service was seen as a good item.

The collaboration of the MAST with HEASRC and Penn State on this project has helped save costs while taking advantage of the expertise available.

KEPLER

MAST had had many issues with the Kepler mission in terms of data delivery. Kepler has changed data format a number of times and in general are not 100% responsive to their archival products and communication with the MAST. MAST has used additional effort in the area to allow Kepler data to be served to the community in an easy and useful form. We are particularly impressed with the speed at which Kepler data appears to the public once delivered to MAST.

The MUG notes that it is very good that MAST is going to be involved in planning the Kepler Legacy Archive. The Kepler team has clearly benefited from feedback that MAST has given them regarding what users want/need, and MAST is in a position to help the legacy archive be all the more valuable.

Good to see that the Kepler team has taken it upon themselves to prepare more useful data from the community. Last year, that was a big issue. This exhibits an interesting role for the MAST; they can't make all the data perfect before they provide it to users, which is beyond their scope. But they can provide feedback to the various missions of data-submitters as to what the community finds useful and what it finds unusable.

The Kepler “colors” project, allowing cross matches of Kepler field of view locations and sources with other catalogues such as GALEX and the Howell-Everett UBV survey will be a wonderful service for users wishing to make use of archival Kepler data or propose new Kepler observations. One suggestion from the MUG is to have a service here to allow a “finding chart” to be produced for a selected region and show the image data available for any survey with an overlapping footprint.

The co-trending basis vectors (essentially a way to allow users to perform ensemble differential photometry on their light curves) are a good product to serve. There is no need to develop this service further, as the Kepler mission itself will be performing this task as a part of their pipeline and have all the light curves re-run by July 2012.

The Kepler FFI's offer a useful community product, yet to be fully utilized. We suggest tool development here as well in terms of allowing a user to see and download cutout regions of the FFI's and perhaps even provide a rough “light curve” covering all the FFI's under consideration.

HST

Work on the HST OTF pipelines has been a great improvement here. It is nice that the new LINUX server really has had the order-of-magnitude impact that was anticipated. Very well implemented! MAST has done a good job working with the HST instrument teams on the ACS CTE issue. MAST has worked to come up with an approach that gets the users the data they need, while at the same time not overwhelming the archive with oversized files, or over-long reprocessing times. This development was not something they could have possibly planned for, but they have worked well about how best to adapt the archive products.

JWST

While a bit into the future for flight data, it is of value to begin the archive of test data for the instruments as both of potential user value as well as test cases of flight data. The Python based coding seems appropriate here and the MUG commends the study and final selection of the CONDOR software. The use of this by a number of current/new large programs such as NOAO (LSST) and Fermi lab (DECCAM) should allow cost savings and forums for collaborative projects. Reuse of the MAST and HST archive software is of value for cost savings and a plan to have the archive ~90% done before launch is well thought out.

In general, MAST is dealing well with the upcoming JWST ramp up. For the next few years, all the JWST-related efforts will be done under JWST funding, so MAST won't have to spend any of its own money on JWST, and they wisely left that out of their NASA contract. It was good that they saw the possible overlap and made decisions about who should support what.

MAST & VAO

The MUG sees the VAO work and continuing on a good path. The archive has a nice interface for such use and we commend the desire to combine with MAST search tools. The MUG suggests that MAST look at ways that will not duplicate effort such as their idea of using the HLA cart.

The web-based search tools provide a nice visual interface for the archive data products. We assume the classic and web-based interfaces will work toward a merger and not continue to develop in two completely separate endeavors.

User Survey

The MAST annual user survey was taken to heart this year with 10,000 surveys sent out; the number of replies being near 400, far in excess of past years. Good news from the survey results, as they gave a positive response to most MAST services and products with the many "not sure" answers related to new or

unfamiliar services. To get survey results throughout the year and perhaps from a more direct user community, the MUG suggests that the MAST look into putting survey links on help-desk calls and archive downloads in addition to those sent out during the yearly survey window

User comments from people who had used CasJobs were generally positive. We liked the idea to show the full CasJobs-type database query to the user whenever any query is made, so that users can be encouraged to learn and experiment.

Public presence / EPO Activities

We realize the time it takes to prepare for and attend meetings or to hold workshops. MAST staff has done a good job in this area and it appears that they plan to continue their effort. We suggest they aim to highlight new products and new mission data and tools they have.