

**REPORT OF THE MULTI-MISSION ARCHIVE AT SPACE TELESCOPE
(MAST) USERS GROUP: JULY 2008 MEETING**

September 2008

The MAST Users Group (MUG) met at the Space Telescope Science Institute in Baltimore on 18 July 2008. Present were committee members Leonidas Moustakas (Chair, JPL/Caltech), Christine Chen (STScI), Duilia de Mello (GSFC/JHU), Steve Howell (NOAO), Casey Papovich (Texas A&M), and Richard Wade (Penn State University).

The MUG was uniformly impressed by the resources MAST has developed, many of which have come a long way in just the past couple of years. The *Hubble* Legacy Archive (HLA) is a particular stand-out. The greatest and persistent critique overall is with respect to the need for greater exposure of everything that is available to the community at large. In the following, we cover the main issues that were identified, and provide feedback and comments.

The value of MAST

The MAST capabilities that allow users to identify, browse, and preview datasets has improved vastly in the past year. MAST has made a number of good interface improvements and revisions to handbooks. We applaud these efforts, and believe that MAST is rapidly approaching the status of ADS: it is not possible to imagine doing astronomy without it.

Exposure of MAST resources and publicity

While we appreciate that each year, the MAST scientists make an ever more forceful effort to advertise the available tools and new developments, we find that it is still not nearly enough. This is some part due to MAST's success: tools and resources are continuously evolving and becoming more sophisticated. This should be an opportunity, not a drawback, however. We urge a renewed push to advertise what MAST is capable of, and what features are in development, to attempt to ensure that it is on scientists' horizon. In addition to the MAST/*HST* Newsletters, there could be frequent updates through the AAS Newsletter. A strong presence at a variety of professional meetings (not only the AAS) is recommended, possibly with dedicated demonstration and discussion sessions. Could some coordination be pursued with astronomical education-research resources? Finally, we highly recommend that MAST be present at the 2009 IAU General Assembly in Brazil.

It is also the case that these tools have become so feature-laden that a live demonstration of their capabilities (such as was given to the Users Group) is almost the only efficient way to learn how to use them effectively. This emphasizes again the importance of outreach to the user community – either via well-constructed tutorials or in face-to-face sessions. MAST should not be shy about using resources explicitly for training possible users.

Planning for bandwidth

Additional exposure, ever-larger datasets, and the increased capability in identifying and re-

questing them, will have its logistical consequences in demands for greater bandwidth. Based on what was reviewed in 2008, we strongly support any and all efforts to anticipate bandwidth needs as pro-actively as possible. We urge the relevant funding agencies that this is an important facet of future MAST functionality and success, and therefore critically important for the entire community’s continuing astronomical research.

External datasets, standards, and NVO

There are persistent choices to be made between in-house infrastructure for organizing and delivering data, and NVO-derived standards. As stated in earlier reports, we support the plan to adopt NVO standards for newly ingested datasets, while continuing MAST-specific developments. Overlap between MAST and NVO activities are clearly inevitable, but the MAST focus should continue to be on MAST resources. MAST should strive to adhere to the NVO standards (as they are doing) yet not develop or duplicate efforts funded by the NVO project. We believe that MAST is striking the correct balance between MAST and VO activities.

As an archival data server, MAST must insist that the datasets it accepts responsibility for (in particular High Level Science Products) are appropriately documented and stable, so that the end user knows what they will get. Although we have not yet seen it, we expect that the FUSE archive manual will provide a good example to follow. In general, FUSE seems to be taking the right approach toward closing out a mission.

We recommend MAST to continue cross correlations with large non-NASA derived ground-based surveys (e.g. SDSS and Pan-STARRS). If required, a modest effort could be made to ensure that these and future datasets are sufficiently NVO compliant for ingesting, but we do not recommend any significant resources be dedicated to developing new independent tools for ingesting or serving such non-funded datasets. At the very least, the committee suggests that such datasets should meet a specific level of documentation before being ingested by MAST. Perhaps a temporary ad hoc committee could be set up for each new mission, to review the archived data and the user documentation available. FUSE, for example, is doing the correct thing by having by have final documents available in detail.

In this vein, uniformly high quality documentation is still lacking from many datasets, e.g. from EPOXI and GALEX. In cases where the onus of good documentation lies with the institutions/PIs that provide data to MAST, it would be helpful for MAST to pro-actively make available guidelines or a check-list of documentation information that is required along with contributed data.

It is reasonable for different NASA-funded archive efforts to push development in parallel, up to a point: synergy is good, needless competition for scarce resources or duplication of effort is not. MAST should continue to be aware of efforts elsewhere, so as to concentrate its own efforts on essential tasks. It is not clear, for example, that MAST is especially well positioned to undertake development of time-domain analysis tools, e.g. for the *Kepler* mission, although it should take a leading role (if need be) in defining standards for recording the time coordinate in archival time-

series data. This would make possible e.g. a quick-look ability of time-domain data.

The Hubble Legacy Archive

We cannot help but bring special attention to the *HLA* as with last year, and how impressed we have been over its deployment and simple functionality and power. It is a job extremely well designed and executed, and a gem for MAST. The emphasis of the *HLA* is, naturally, on *Hubble* data, and this should continue to be the case. The model is so good, however, that we encourage MAST to keep open the option of ingesting non-*Hubble* data in the future, if appropriate resources, funds, and time become available.

GALEX

We tested GALEXView and were impressed with the quality of the tool and congratulate MAST for the improvement in the GALEX search page since last year. GALEXView is a good interface and offers a nice yet powerful web tool for archive users to view object information in context. The use of this type of tool for other archive products could be attractive.

However, we still ask MAST to provide more and better documentation throughout. For example, the images and photometric results are not easy to interpret. (“What is the aperture size used for the ”magnitudes” listed?” “How are ‘counts’ converted to magnitude?”). For example the ERO and GR1 HTML documentation pages are off-site at Caltech rather than in-house, and stale (including, even, broken links).

We also recommend that users have the option of saving the results provided when double clicking in objects within the GALEXView screen. At the time we tested it, a pop-up screen appears on the left side, but users have to write down the numbers. If saving a table is not possible or too complex to be generated, the ability to copy and paste numbers directly from the pop-up screen to another editor would be a fix.

We note that these tools, as they become more powerful, also become more complex to use, and in some cases may represent “overkill.” In GALEXView, large amounts of data are transmitted to the client machine, which needs to be correspondingly powerful to handle the flood. Not every archive retrieval task needs to be so graphically oriented or bandwidth intensive, and we urge that simple means for simple tasks be maintained.

Other

It would be useful to know the percentage of MAST data which were never published and develop tools to stimulate users to work with these data.

Work on spectral products seems somewhat behind imaging efforts. In particular, the spectral containers work appears to be leading the world effort and should push a bit to get definitions of Echelle, polarimetric, and spectral time series data defined. The “spectral classes” project seems a small PI-level effort and the committee was not clear on its purpose and ultimate goal.

In cases where spectroscopic data are available (e.g. for perusal), we recommend MAST to also provide these data in 1-D FITS format or as plaintext tables.

Color DSS images and the connection to Google Sky are clear directions to go to enhance public and other visibility for MAST but large effort in this area should be tempered, against highly useful scientific products for serious archive users. We do not discourage color and highly visible products, but caution against going too far in this direction. We again believe that MAST has found the correct balance in this area.