MAST Users Group Report – July 2010

The MAST users group (MUG) met at the Space Telescope Science Institute on 16 July 2010. MUG members are Mike Crenshaw (Georgia State University), Steve Howell (Chair, NOAO), Casey Papovich (Texas A&M), Evgenya Shkolnik (DTM), Ben Williams (University of Washington), and Jay Andersen (STSci) and all were present at the meeting. This was the first MUG meeting for Andersen and the third for Howell and Papovich, who agreed to remain on the MUG for another year.

Overall Comments

The MUG was impressed with the work undertaken and completed by the MAST group during the past year. This was especially impressive given the delay of nearly 9 months in NASA funding for the archive group mission. 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.

Power of the Archive

It is clear that the archive for a mission is truly a force-multiplier in terms of its impact. Thus, it is very important that the archive offer easy access and provide useful data. As an example, it appears that the Kepler mission was very unprepared and not aligned with this well known ideal.
The MAST archive folks are doing a heroic job helping that mission and turning their data into files and data products that people can actually use, both the full-field images and the time-series. It is good that they're already far into thinking about JWST and how they will best be able to serve-up that data.

It is very clear that the HLSP developed and served by MAST have a much higher distributed data volume compared to data archived and compared to the normal data products served for the supported missions. The HLSP have a much higher archival demand and utility to the community.

As in the past, more than 2/3 of HST publications use archival data (1/3 use GO (PI) data only). This is yet another clear indicator that MAST and good archiving provide strong returns on science for a fraction of the mission cost. Other missions should be required to consider and develop their archive infrastructure as the missions are developed.

When archives win contracts to serve the data, they should take into account the Kepler-lessons learned and determine ahead of time whose job it is to put the data into a useful form. The original mission proposal should include information

about how data products would be brought to the public archives, MAST should be involved, and NASA should require and monitor these activities.

It is clear that MAST is not the only player on the block for data archiving and other NASA centers would clearly like to take some of the "market share", by providing some enhanced products. The MUG particularly appreciates MAST's interest in doing the "little things" (such as adding WCSs to Kepler footprints) that make things infinitely more useful to more users.

HLA/NEW INTERFACES

Some of the "gee-whiz" things developed for the HLA may initially appear to be more for the general public than for "serious" astronomers, but it is clear that if the products can be put into directly usable form, so that scientists wouldn't have to either photo-analyze jpegs or do their own mappings to the sky, they can use this high-level data just like they use SIMBAD or any other database full of information. It seems that MAST is actively crossing that threshold now, and it is good to see it complete the overlaps between the functionality of the HLA, and the other missions and tools such as the VAO, APT, STARVIEW, etc.

The new interfaces for MAST seem to be the way to go. Users previously had to make their own "footprints" to show what could be "poached" in the archive but it is great that MAST is making these tools easy for everyone to generate. GALEX astroview is an interesting and powerful browser interface, like WWT or Google Sky. Could this be used for HLA viewing in the portal? Need to be sure there is not duplication of effort here.

As for the new interface, it will be important to keep the low-level (single-image) data available. It could be that in the new interface it will be easier for users to see "all" the images. Currently, MAST only reports "associations", which often contain many exposures, and it's impossible to know how many individual exposures there are (shorts? deeps?) without either downloading them or looking at the phase-2. Perhaps with the new format, they will include this flexibility.

It would be useful to have some kind of a "depth" map, that shows not just how many exposures there are at a point, but how deep it goes. Some kind of useful units that properly account for the different efficiencies of WFPC2/ACS/UVIS/etc may also be useful.

Supported missions – specific comments

One of MAST's current goals is to produce a common data model (and database infrastructure) for all data from all missions to simplify organization. HST

operations is supporting this with new hardware and databases in OPUS and DADS. This is good! Improved efficiency and simplicity and places much commonality and mission footprints into HLA footprint database. The MUG strongly supports this vision.

The Kepler work has been great and has allowed archive/GO use at a high level. MAST's work to make new scripts and products for light curves using Kepler data is sound and will be useful. The MUG believes that it is important to have someone intimate with a mission (e.g., Kepler) on the MUG during the early stages of the archival work and the current chair serves this role for Kepler.

The MUG noted that JWST will generate a significant ramp-up for MAST. Are they prepared to meet this ramp up in terms of monetary and people resources?

High Level Products

MAST might explore the idea of sharing, with the community, contributed GO/archival users high-level data products. This could be a force-multiplier similar to the archive itself but may also carry increased support and use loads.

Perhaps a method by which the user community might learn about high level products would be if MAST tools pointed to high-level products at the same time as they point to refereed publications in the standard column output of the archive. These links in the archive searches may make the public gradually aware of this powerful way to matrix data products. Perhaps
there could just be a check-box for "high-level products" that would be separate from INSTRUMENTS and SPECTROGRAPHS...

The MUG was very impressed with the clearly useful high level products and the new directions the MAST plans to follow. Data release 5 will be wonderful. The MUG also liked the interactive tools being developed and the "one box" interface.

The MUG still looks foreword to the PASP paper describing the details of the many new and wonderful features.

CasJobs

Point-source matching for CasJobs calls that query the databases for GALEX, ISO, FUSE, etc. is of high interest and the MUG is aware of the "issues" here in terms of various missions pointing and catalog location precisions. We are certain MAST will resolve these issues in an elegant manner.

Is MAST/HST planning on a point-source catalog of objects detected in its archive images? This large effort could provide an enormous opportunity to

cross-calibrate instruments and evaluate HST's calibration over time. This would add a huge boon to people being able to use the archive directly for science

Using this sort of interface across all missions is a good idea and makes for ease of use for archive scientists. Adding in the ability to point to sources in other mission archives (e.g., Chandra, Spitzer) would be great as well. We realize much of this depends on the non-MAST archives hosted for those missions.

CasJobs sounds like a great way for advanced users to access all MAST mission data using SQL. Documentation, and working examples should be provided for novices.   Some of this documentation seems to exist and we would encourage MAST folks to continue that enterprise.

MAST & VAO

The MUG continues to be happy about the MAST involvement in the VAO. MAST is a world leader in this area and as such provides a large, yet mostly unknown and unsung resource for astronomy. MAST sets an example for other archives and "wanna be" archives to follow. It is very encouraging to see that the funding agencies have (again) decided that the VAO is important and provided long-term funds for such work. We hope it continues.

The VAO plan presented came across as a bit high-level and thus not completely clear as to the exact details. This is likely a result of the fact that VAO funds and ramp-up have just started (again). MAST has a major role here, and there are a frightening number of details that were not likely able to be conveyed in a brief talk.

In a connection to EP/O activities, it does not seem like much of the VAO enterprise is consciously aware of how professionals and non-professionals will make use of the archive. The ease with which APS can be written to serve to the public makes this a very obvious route. "Pretty pictures" to capture attention and hook interest must be balanced with actual scientific results and deeper level information. It is also ok to say what we don't know about things as well as what we do know. Some of the unanswered questions are things that spur people to enter science.

The goals of VAO seem correct (use professional/industry standards for dataset conventions, standardize data access and discovery protocols). However, it seems hard to see how MAST/VAO will be able to enforce this on other VAO contributors. Incentives, such as funding agencies requiring award recipients to submit VAO-compliant datasets as part of their funding requirements, are needed to be put in place.

Data Serving and Bandwidth

The MUG was impressed at how quickly the MAST has been able to address problems that come up in terms of bandwidth or server speed. They are doing a good job focusing on both the present and the future (while serving up the past). They have done a good job transitioning from funding sources (as missions end), and dealing with the loss of expertise and help from the European mirror site. The change from Solaris to linux is applauded.

We agree with the principle that the Institute's access to the internet should not be the bottleneck for the Archive access speed and the Archive should not prevent the Institute from having good access either. It sounds like efforts are underway to ensure this. One possible idea is for MAST to have its own dedicated internet line to the outside world that does not share or compete with that of the Institute. Once they start serving more things from a "static" archive (or at least one that is background updated), it may be easier to have mirror sites.

The MUG is aware of and wanted to make a specific comment on the fact that Outside storage and backup resources are not free. Both in terms of actual monetary cost but in time to MAST as well. Constant work to identify, buy, and use the latest, longer term, best storage devices requires an effort by MAST personnel generally not recognized. Data volume is increasing exponentially as is disk storage size per cost unit, but data access bandwidth to such material is not.

2011 Senior Review

MAST will need to prepare for the 2011 senior review during this upcoming 6 months and will need to weigh tasks to work on and finish vs. senior review presentation materials and demos. The MUG feels that MAST leadership is well aware of this and their experience is the best determiner of this balance.

The "one box" interface, VAO work and plans, and the new high level tools and interfaces are highlights for the review. The level of archive use and interface for the newest missions are valuable to show off as well. MUG members are willing to help in this review at whatever level the MAST feels is required.

User Survey

One way to measure and get feedback for a project is via user surveys. MAST has done this for years and should continue. The MUG feels that MAST may try some things to redouble their efforts to get higher response numbers and to keep track of long-term trends and desires by users. Asking similar questions over time and seeing how the response varies may be a good indicator of interest, or

lack thereof, for specific tools and products. Perhaps add a survey link to every help-desk reply and collect user survey information throughout the year.

One reason for low responses may be that people are happy and satisfied with the services and have nothing to complain about. MAST should not interpret the unreturned surveys and the "don't know" answers as "don't care". The mission-specific user forums may be made more effective with a large obvious link on each search page.

We understand that much of the documentation is not under MAST control, but suggest a MAST specific documentation page which links to the latest and greatest documents.

Public presence / EPO Activities

In working very hard to provide the tools and products MAST users want and expect, MAST has little time to blow their own horns and tell the community, one on one, about their archive. Additional public presence would be helpful in terms of advertisement but as a way to get direct user feedback. It is always good to see a MAST presence at AAS and other major astronomical meetings.

MAST members would be good to send out to the community to give colloquia about the "how to's" as well as highlighting archive use. They should give a wide array of examples of what kinds of science is enabled with enhanced searches.

The MUG encourages an archive workshop with a format similar to a "summer school". We also encourage the various NASA archives to meet regularly to share tools and methods and to discourage duplicate effort.

The MAST EP/O plan at present seems small in scope and of limited impact. We'd like to see an expanded program fueled by additional EP/O funding from NASA.

Visualization tools, such as the iphone application, are great ways to get the public involved and interested. Perhaps always have a way for them to "get more" information and be able to drill down as deep as they wish to go.