MAST Users Group Report
January 2018

The MAST Users Group (MUG) meeting was held 01/23/18–01/24/18 at STScI. The members of the group are Dan Foreman-Mackey (Chair; Flatiron Institute), Tabetha Boyajian (LSU), Martha Boyer (STScI), Knut Olsen (NOAO), John O'Meara (St. Michaels - Vermont), and Dan Weisz (UC Berkeley).

Mission Statement
The meeting started with a discussion of the core goals of MAST and the identification of a possible tension between the goal of enabling cutting edge science and the goal of lowering the barrier of access for the MAST holdings of public datasets. It is clear that the relative importance given to these goals will affect the allocation of resources within MAST. The MUG recognizes that these are both important goals and it is recommended that an official stance be developed for the balance of these goals. This discussion is especially timely because, at present, the community relies on MAST for archival data distribution and GO proprietary data retrieval. For HST in particular, all data will soon be archival, and MAST will be the sole tool to enable the community to do science.

Community visibility
The MUG is impressed by the many initiatives at MAST either nearing completion or initially underway. Nevertheless, the MUG is concerned that many of the great tools coming from MAST are being missed by the community. The MUG feels that more effort should be given to community outreach, and offers the following suggestions:
1. Consider MAST 'hack sessions' at STScI meetings/workshops. Given the significant scientific diversity of the community that visits the institute, small hack sessions or workshops that highlight how MAST can enhance the science discussed at the meeting could have significant impact.
2. Create a small (1 or 2 page) set of slides that can be added to presentations so that speakers who use MAST in their research can highlight that fact
3. Consider establishing a small set of MAST 'evangelizers' who can give presentations to the community. The MUG themselves and their contact base are a good starting point.
4. Update the portal to have a rotating 'MAST highlight' banner that links to the resource.
5. Have MAST staff prepare white papers for the Decadal Survey. MAST represents one of the premier archives for astronomical data, and can be used as a model for other existing or upcoming archives.
6. Prepare a set of short youtube videos highlighting MAST products and their use. Place them on the portal (perhaps in conjunction w/ #4), and on astro social media (Astronomers facebook page, twitter, etc). Link to them in the Call for proposals in the Archival proposal section.
7. Consider policy changes to archival proposal calls that encourage more interaction w/ MAST.

Open Source
The MUG emphasizes the important role that open source software plays in the development, mission, and visibility of the MAST. The MUG encourages continued and increased support and development of open source pipelines, tools, and interfaces for interfacing with the MAST's data products.

NASA Astronomical Virtual Observatories (NAVO)
The MUG was given an update of NASA participation in the Astronomical Virtual Observatory. The MUG reaffirmed its view that NASA's continued participation in the VO is important and a strategically prudent decision that enables the US space-based astronomical community to draw from a wide range effort and expertise as it designs its data archives and platforms. The development of a registered TAP service for NASA archive databases is a useful step, particularly for the development of Science Platforms that use the TAP protocol and exposure of MAST holdings to TAP clients such as TOPCAT. A strength of the VO lies in providing widely agreed-upon protocols and standards for individual data centers to implement as they develop their data facilities and tools, and the NAVO effort is contributing to this. The MUG was pleased to see effort directed toward making the breadth of the VO-effort effort more visible and useful to the community through its advertising campaign and its planned integration with the AstroPy project. The priorities presented for the NAVO effort in 2018 are well aligned with the MAST mission of enabling science with its data holdings.

Science Platforms
The MUG was introduced to the "science platforms" or "server-side analytics" development being investigated at MAST. A major component of this is the development of a technology stack for bringing calculations to the data. The MUG recognizes the importance of this work in the era of growing datasets and it is clear that this has the potential to significantly impact the scientific workflow. This could enable novel analysis of large datasets by a small number of power users, but (perhaps more interestingly) this could also lower the barrier to entry for new researchers in the astronomical community and beyond. This could have an especially large impact on the workflow for analysis of upcoming missions like WFIRST with a fast release cycle and large data volumes.

The MUG suspects that the development of the technology stack will not have the intended impact without a specific target user base in mind and without parallel development of comprehensive pedagogical resources. The MUG recommends that a set of "user stories" be developed to clearly demonstrate the target user base for these platforms. Some inspiration may come from the tutorials that accompany the APT proposal preparation tool.

Notebook Archive & "Hand-off"
There was some discussion of a workflow where the MAST documentation/tutorial pages could include a "hand-off" button allowing the document to be downloaded as an interactive Jupyter notebook or launched on the science platform system. This could substantially increase the usability of the MAST resources and lower the barrier to entry. The MUG strongly endorses further development of this tool. The MUG envisions a time in the future when the "download

basket" button in the portal is replaced by a "generate notebook" button.   The issue of notebooks as a stable future platform came into question (along with Python as the dominant language), but these concerns are minor, especially in consideration of the wide adoption in the broader community.

MAST API & AstroQuery interface
A programmatic interface (API) to the MAST data products was presented. This interface conforms with REST standards and it exists as both a set of well documented HTTP endpoints and a Python client library within the AstroPy affiliated package Astroquery. This feature has been recommended as part of previous MUG reports and the current MUG was thoroughly impressed by the progress and the details of the implementation. Members of the MUG who had no previous experience with APIs or Astroquery were given an opportunity to use the Python interface and were all able to construct queries related to their research interests within minutes. The main limitation of the current API seems to be a lack of queryable mission-specific columns. The MUG encourages continued support and development of these interfaces with an emphasis on documentation and tutorials.

High-Level Science Products (HLSPs)
There was a discussion of the impact and scalability of the current high-level science product (HLSP) workflow. It is clear to the MUG that the HLSPs are a useful resource and with a huge impact. While the current HLSP ingestion workflow is hands-on, the MUG believes that these are sufficiently important resources that continued development and support should be a priority.

Discovery Portal & Website Redesign
The presentation to the MUG did not include much information on the status of and plans for the MAST Portal. Since the Portal is the main MAST interface for many users, the MUG recommends that the Portal be a focus in the next year, either to improve it or to replace it. The MUG believes that the recently hired UX/UI expert within MAST will substantially help this process and recommends this focus on user experience continue. The MUG was also pleased to see the progress on the website redesign efforts and expect that this will significantly reduce the barrier to access of the MAST resources.

Kepler/K2 Closeout & Planning for TESS
As the Kepler/K2 mission closes out, MAST is transitioning to becoming the primary resource for Kepler/K2 archival data access and support. Furthermore, MAST will be the archive for TESS data (and, in particular, the TESS full-frame imaging) within the next year. These large time series datasets present novel challenges for MAST and some of the solutions being pursued by the MAST were presented. One such solution is a "cutout service" for the TESS full-frame images where a user will be able to request to download a small patch of the imaging instead of directly accessing the full dataset. It is clear to the MUG that this is a crucial interface that will enable easy access to the TESS data for users with all levels of experience. The MUG believes that the MAST is well equipped to provide this service and wholeheartedly encourages its continued development.

DOI Project

Tracking datasets through DOIs provides traceability and permanent online access. MUG commends MAST for its efforts to encourage users to create DOIs associated with their publications. In 2019, MAST will work with AAS Journals to add 18 institutions to program that queries authors to provide DOIs during paper submission and the MUG recommends expanding this program to more institutions and to other journals. A partnership with ADS might also be beneficial; a link to data DOIs from a paper's main ADS page would be one way to add a DOI after a paper is published. The plan to make DOIs "mandatory" for JWST publications by tracking arXiv submissions is a good strategy that works for ALMA. MAST might consider a trial period for Hubble publications.

Astrometric Improvements for Hubble

The MUG is impressed with the substantial astrometric updates coming to Hubble data in MAST in the coming year. The improvement achieved through use of the Gaia data will be critical for science relying on accurate astrometry (e.g., proper motions) and for planning JWST observations.

Contributed Services & ExoCTK

The "Exoplanet Characterization Toolkit (ExoCTK)" was presented as an example of a possible "contributed service" that could be used to provide research area-specific interfaces to the MAST holdings. The MUG was concerned by the potential technical debt that could be accrued by the MAST's offer to support services that are not directly developed by the MAST team. The MUG does not recommend this as a focus of the MAST, but instead suggests that some research be done to develop a clear description or "user story" of the potential use cases.