

## REPORT OF THE MULTI-MISSION ARCHIVE AT SPACE TELESCOPE USERS GROUP (MUG): OCTOBER 2001 MEETING

The MAST Users Group met at the Space Telescope Science Institute in Baltimore 8 October, 2001. Present were Committee members Luciana Bianchi (JHU), Jane Charlton (Penn State), Andy Fruchter (STScI), Bob Nichol (CMU), and Paula Szkody (Washington). Thomas Ayres (chair, Colorado) participated from Boulder via videoconference.

This was the first meeting since the recent inception of the MUG, and the primary purpose was to acquaint the membership with the functions of the MAST and the key issues facing it. The meeting consisted of a number of presentations by MAST personnel concerning the internal workings of the center, its various current and future projects, and the external context within which the MAST operates. The MUG in general was very supportive of the current accomplishments and near term directions of the MAST, particularly the proactive stance of the MAST leadership in identifying and facing future challenges. Although the main purpose of the first meeting of the MUG was informational, the committee does wish to highlight the following findings and recommendations:

- **Sloan Digital Sky Survey** We endorse the MAST's efforts to archive a copy of the SDSS and help develop an interface to enable outside non expert users to gain easy access to the vast array of imaging, photometric, and spectroscopic data contained in the quarter-sky survey. We view the SDSS both as a key astronomical resource as well as a model of the large databases that are becoming commonplace in observational astronomy. MAST's stewardship of the SDSS and the facilitation of easy access would be important steps toward the National Virtual Observatory. Significant new resources will have to be identified, however, to permit the MAST to take on the SDSS activity.
- **Spectral Scrapbook** The Spectral Scrapbook browse facility allows ready access to spectral information concerning a wide variety of celestial sources collected by a number of NASA observatories. The Scrapbook provides an easy way to visualize the nature and quality of selected observations of each target in its database. The committee felt that the Scrapbook should emphasize access to all of the relevant spectral observations of a target, in addition to the small representative "reconnaissance" set. Then, the Scrapbook could serve as a direct portal into the archive in addition to its overview function. This might be accomplished simply by highlighting existing links. A second priority should be the development of manipulation tools such as would allow a wide wavelength range of multispectral data to be displayed in the same window.
- **STARVIEW** STARVIEW is one of the original query tools developed to access *HST* archival data, and now expanded to include the other datasets under

MAST's purview. STARVIEW is somewhat more complex to use than the basic MAST Web query interface, but allows for more detailed searches. In addition, one of STARVIEW's more popular options is the Visual Target Tuner (VTT), often used in proposal planning but also in post-facto comparisons, say, of science images with the corresponding fields from the Digital Sky Survey. One drawback identified by the committee is that Java-based STARVIEW is platform dependent, and requires installation and maintenance on the local users system. The committee felt that migrating STARVIEW to the Web would solve both problems. Furthermore, the committee proposed that the STARVIEW functionality (and that of the Scrapbook as well) could be subsumed within a comprehensive query interface so that the user would be confronted with just a single portal into the MAST, not several.

- **Astrophysics Data Centers Coordinating Committee Recommendations** The ADCCC, a coordinating committee for the various astrophysics data centers, has recommended that NASA reconstitute an oversight committee, something like the disbanded SOMOG, which would be able to examine and set priorities within the distributed system, and develop policy such as system-wide data standards (e.g., defining world coordinate systems). The MUG concurs that such an oversight committee would be valuable; indeed, perhaps a necessary step on the road toward the creation of a National Virtual Observatory. A second, key, ADCCC recommendation involves the role that the centers should play in helping new missions with data processing and archiving activities, and is described in more detail later.
- **Insularity** The challenges facing the MAST, and the other astrophysics data centers, are not unique within the broader context of NASA (and NSF) research. The Planetary Data System and the Earth Observing System, for example, are responsible for managing and distributing massive volumes of data to their users, and both the space and groundbased solar communities are intimately familiar with archiving and accessing substantial complex databases. The NSF-supported SOLIS project (synoptic monitoring of the Sun and its magnetic activity) alone will be accumulating *terabytes* of information *each day* when it becomes fully operational next year. Of course, a great deal of thinking on the issues of data storage and management go on outside the confines of astronomy and the geosciences: Computer sciences departments in the Universities are bristling with expertise in these and related areas. The MUG is concerned that the MAST in particular, and the astrophysics data centers in general, have not been taking full advantage of the knowledge and experience accumulated in these parallel efforts, and that unnecessary resources might then be devoted to essentially reinventing the proverbial wheel. More communication with, and sharing of information between, the various archival and processing efforts NASA-wide would, in the view of the MUG, be quite beneficial. Breaking

down the cultural and technical barriers between these historically independent efforts will not be trivial, but the potential rewards are great. Along these lines, the MUG recommends that the MAST take advantage of initiatives such as the “Technical Seminar Series” at STScI, and expand them to draw upon interdisciplinary expertise from the Universities (and perhaps Industry as well). In addition, the MUG suggested that an increased sensitivity to the growing numbers of LINUX users would be beneficial.

- **Future NASA Missions** Upcoming NASA UV/Optical missions such as GALEX, or those in the planning stages like PRIME and KEPLER, promise to collect massive amounts of celestial data, which will be a challenge to store and distribute in a public archive. According to the ADCCC recommendation, it is essential that the appropriate archival centers, such as MAST for UV/Optical missions, be involved early-on with the planning for data processing and archiving and that sufficient mission funds be set aside for that purpose. Oftentimes, mission teams will skimp on the analysis side in favor of the hardware effort, creating a difficult situation down the road when the mission is active, but processing and archival tools are not ready. The MUG also strongly endorses the ADCCC concept that the archival centers serve a role as data processing and storage facilitators for the new missions.
- **Future of Archive Services at Space Telescope** FASST is an internal effort at Space Telescope to identify the future needs of archival data users, so that the corresponding hardware and software issues can be anticipated in advance. Part of the effort is to analyze how astronomers “do their science,” and develop strategies for archival structures and search engines to facilitate that activity through improved “data discovery” and access. For example, at present it is difficult to use the results of one search to initiate another. Improvements are needed in metadata, and in world coordinate systems. FASST also is overseeing the transition of *HST* instrument data from active to heritage mode, particularly to ensure that proper “end-game” planning is taken by the particular instrument teams prior to decommissioning of their instrument and dispersal of the team’s expertise. FASST is related to the NVO effort in the sense that coordinated catalog services (“one-stop shopping”) will likely be an important model for NVO query strategies. The committee views the FASST activities as an example of a proactive response to the future, and strongly encourages such efforts.