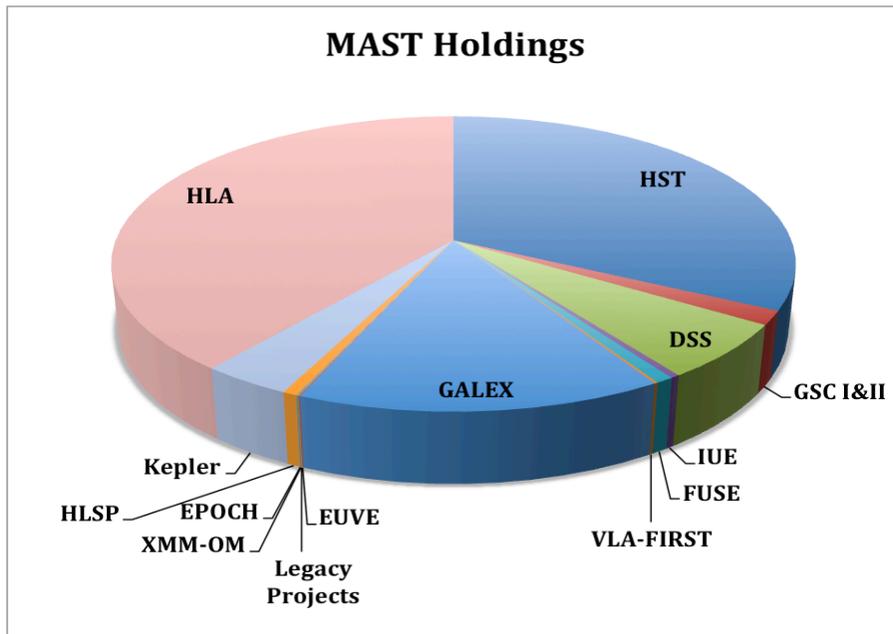


NASA Data Center Annual Program Plan
NASA Grant Number NNX09AF08G / STScI Grant J1346
March 2010 through February 2011
The Multi-Mission Archive at Space Telescope Science Institute (MAST) Phase II
(Optical/UV Science Archive Research Center)
Space Telescope Science Institute
3700 San Martin Drive
Baltimore, MD 21218

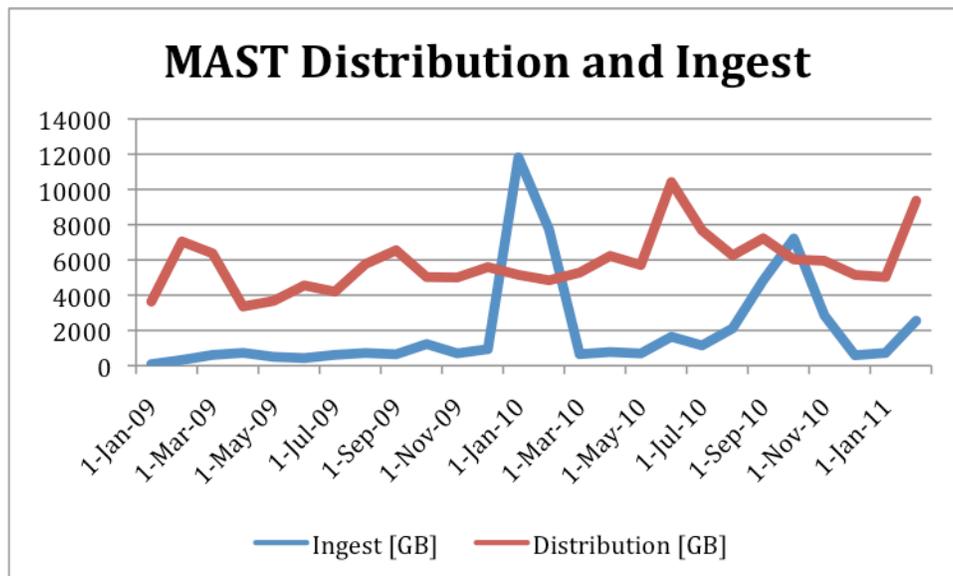
Overall Mission: MAST supports active and legacy mission datasets and related catalogs and surveys, focusing primarily on data in the ultraviolet, optical, and near-IR spectral regions. Support includes providing data curation, providing expert support to users of the data, providing access to data-specific calibration and analysis software, providing user support for this software, and maintaining public access interfaces to the data. This report covers data financially supported under the MAST contract. Archive and distribution activities for HST data are supported under the HST contract; the Kepler contract supports some of the archive activities for Kepler data. Some of the statistics include HST and Kepler data volume and usage statistics.

Holdings and distribution

As of March 1, 2011 MAST holdings are over 125 TB, including over 64 TB of Hubble Legacy Archive data, 53 TB of HST standard pipeline products and nearly 1 TB of Community Contributed High Level Science Products. The figure below shows MAST holdings as of March 1, 2010.



The archives at STScI have distributed over 48 TB of data between March 2009 and February 2010, and have ingested over 5 TB of data. The figure below shows the statistics on data ingest and distribution to the public from Jan 2009 through February 2011. The large spike of ingest in January 2010 is due to ingest of new HLA products. The spike in the fall of 2010 is due to ingest of GALEX GR6 data.



Mission/Project reports

Hubble Legacy Archive (HLA)

HLA had its fourth data release (DR4) in March of 2010. In this release, reprocessed NICMOS data, reprocessed WFPC2 data (and source lists), new HLSP, additional ACS multi-visit mosaics and updated ACS combined images were added to the HLA data holdings. This brought the total data volume to approximately 64 TB. DR4 also included numerous interface enhancements: improved plotting tool, COS and WFC3 footprints, DADS data retrieval via HLA cart, Inventory table enhancements and a consolidated Help Center.

Following DR4, work started for DR5 planned for release March 2011. This release focused on interface enhancements primarily a new Footprint Interface. The new footprint interface capability includes: zooming and panning, multi-instrument selection and full table and cart integration. Other interface changes include: access to multiband source lists, improved download flow from cart, sources clickable in interactive display among others. New data in DR5 include new HLSP, completed WFPC2 processing (over 300 new visits with source lists) and ECF-generated ACS and NICMOS extracted GRISM data hosted at STScI. There will be incremental data releases following DR5 and

will include enhanced WFC3 images and source lists, additional ACS mosaics and prototype COS and STIS products. Release of DR5 is scheduled for March 8, 2011.

Galaxy Explorer (GALEX)

Data Deliveries:

All previous imaging surveys were expanded and released in the form of GALEX Release 6 (GR6). The areal coverage of medium imaging survey and reprocessed GI data in the GR6 represent increases of more than 60%. The AIS was essentially completed with the GR4, but the GR6 represents a substantial increase in exposure over the GR4. Self- and help documentation was updated to reflect the changes in this release. This will be the last major release before the final data reprocessing of the mission.

MAST also received, ingested and distributed 7 Guest Investigator (GI) deliveries during the reporting period, consisting entirely of data obtained through the GALEX NUV camera. Caltech processed all data observed after May 4, 2010 with a new processing pipeline system. The most important change for users was in the adopted magnitude zeropoints of a few percent.

Catalogs:

At the close of the reporting period MAST was in the process of transferring the GALEX-SDSS (Sloan) cross-matching catalogs created by Luciana Bianchi of Johns Hopkins University. These catalogs include the largest attempt yet of differentiating GALEX sources into identifiable astronomical classes, and in the case of objects within the SDSS sky coverage, of doing so with additional input of optical colors (and sometime spectral types) as well.

Tools:

New data fields were added to the results page of MAST's GALEX data browsing and download facility, galexView v1.4 in response to user input. Work was continued to advance galexView to v2.0. New components were built for this tool, including "Astroview," an all-sky imaging viewing tool, and an application for accessing galexView from mobile devices.

Techniques to create a mosaic of the GALEX sky (low and intermediate galactic latitudes) were perfected and implemented for all imaging surveys and image resolutions for GR4/GR5 NUV data. New techniques were identified to create "TOAST" projection images, and by the end of the reporting period the TOASTing of the AIS GR6 survey data was underway.

GALEX Close out:

MAST is participating in a series of meetings to identify products and documentation to be archived from the Project with mission termination at the end of FY11, including a final reprocessing of all data.

Kepler

The Kepler Data Management Center (DMC), funded directly from the Kepler Project, was established at STScI to archive the Kepler Data products. Kepler project team members, Kepler DMC staff, and MAST staff members continued to collaborate to design and implement several search interfaces and website documentation and content. During the past year, MAST implemented some alternative distribution methods for both the team users and for access to the public data. Each quarter, the proprietary data is bundled by program into tar files that are made available via password-protected ftp. The public data are bundled by quarter and are made available via ftp or through the browser. All public data are also available as individual light curves. Ninety-five percent of the Kepler data are distributed via the tarfiles and on-line public lightcurves.

Early in the reporting period the Kepler Project made available a list of dropped targets from its initial list of 156,000 reserved exoplanetary list. It also changed its policy for GOs to propose on any object in the Kepler field as long as the proposed science did not overlap the key mission objectives (discovery and characterization of exoplanets). In addition, the mission provided a list of "false positives" (light curves showing planetary transit features, which upon further review are probably not due to exo-planets), and a list of Kepler Objects of Interest. Because these lists facilitated demand by the community for non-exo-planetary data, MAST places links to these data clearly on its home page. MAST also posts new FFI files on a monthly basis, and these data are used as data sources for proposals in both the Kepler and ADAP programs.

MAST has continued to expand and update its FAQs for the mission and data properties. These are not consolidated in another single website. A MAST staff member participated in the Kepler GO proposal review process.

The KIC catalog was delivered to MAST without ultraviolet color information that allows users to differentiate OB stars from other astronomical sources. To help redress this loss, MAST has cross matched the GALEX GR6 imaging survey catalogs (containing near- and far-UV magnitudes) against the Kepler Input Catalog (KIC) objects in the Kepler Field of View (FOV). MAST constructed two catalogs of 'safe' and 'likely' matches for users and makes them available in three ways: as tables downloadable by ftp, on a "MAST style" request form, and via the CasJobs tool.

The KIC indexes its objects by a unique identifying number. This index is not yet well cross-referenced with identifiers of other source catalogs. Consequently, MAST has cross-referenced these identifiers with identifiers in the infrared 2MASS catalog and made this information available in its target and data search form pages. Plans are underway to add USNO Catalog B identifiers to the Kepler search results pages. The

USNO catalog exercise will expand the number of searchable objects by including many objects in that catalog that were not included in the KIC. This addition will provide completeness of objects within the Kepler FOV down to the USNO Catalog's confusion limit of 21st magnitude.

Epoxi/EPOCH

In 2010 MAST added a new mission, EPOCH (Extrasolar Planet Observation and Characterization), to our portfolio. EPOCH is a NASA-funded mission that used the Deep Impact spacecraft to monitor a small number of stars with known transiting giant planets to characterize those planets and to search for others. EPOCH also measured light curves for the full Earth disk; those data can be used to assess the effectiveness of similar observations for extrasolar planets. MAST is hosting the EPOCH data for the public. There has been considerable interest in these data: since their release in May 2010, more than 200 GB have been downloaded (4 times the total volume of EPOCH data).

Community interaction

Survey

In June 2010 MAST administered what has become a yearly survey to gather feedback about our service and to gauge priorities for future work. There were respondents to the survey. The results and many of the comments have been posted on the MAST website. (<http://archive.stsci.edu/surveyresults/2010/index.html>)

MAST Users Group

The MAST Users Group (MUG) met in July 2010. The MUG provides an essential user perspective on archive operations and development. All the presentations have been posted on the MAST website. The MUG report is posted at (<http://archive.stsci.edu/mug/index.html>)

Outreach

MAST and the HST Office of Public Outreach (OPO) began the work required to make select HST Press Release image available via VO, collaborating to integrate AVM standard meta-data tags into the press-release images and to create a database and associated VO web services. This work is well underway with the database design and ingests software essentially complete. OPO has begun to deliver final versions of the images and meta-data. Creating VO services is the final step. The development work is expected to be complete by the end of March 2011 and this project will then transition to a routine update mode.

Virtual Astronomical Observatory (VAO)

A new employee was hired for the VAO Education and Public Outreach (EPO) effort. The year one program plan is developed and is being implemented. The VAO connections to the Education Standards have been completed, as has a review of other VO EPO efforts. The MAST VAO/EPO representative will participate in the Astrophysics Forum, which has overlapping and distributed collaboration with the VAO EPO work.

MAST staff members have created a “reference implementation” of a Table Access Protocol (TAP) service. Some registry

Other Major work efforts

MAST staff worked on many projects during the past 12 months that introduced new or enhanced capabilities or attributes. We describe a few highlights below.

Hardware and Migration

The hardware purchased in the previous year has been fully implemented. Both Linux and Windows machines have been clustered to improve reliability. Data have been fully migrated to SANs eliminating the previous heterogeneous and unreliable storage situations.

Community-Contributed High-Level Science Products (HLSP)

Eight sets of community contributed reduced science ready data sets were ingested into MAST over the past year. Ten additional sets are in various stages of preparation and ingest and initial contact with PIs have been made for several others. HLSP utilizing HST image data have been included in the HLA. The High-Level Science Products are extremely popular, which is attributable to their science-ready data quality.

