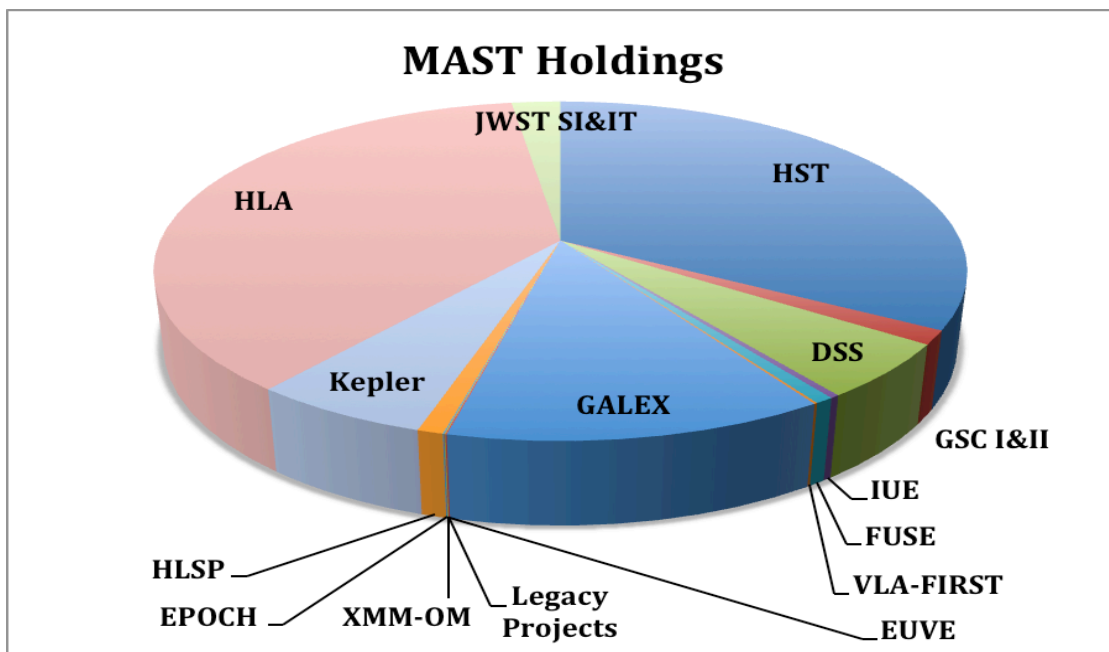


NASA Data Center Annual Program Plan
NASA Grant Number NNX09AF08G / STScI Grant J1346
March 2011 through March 2012
The Multi-Mission Archive at Space Telescope Science Institute (MAST) Phase II
(Optical/UV Science Archive Research Center)
Space Telescope Science Institute
3700 San Martin Drive
Baltimore, MD 21218

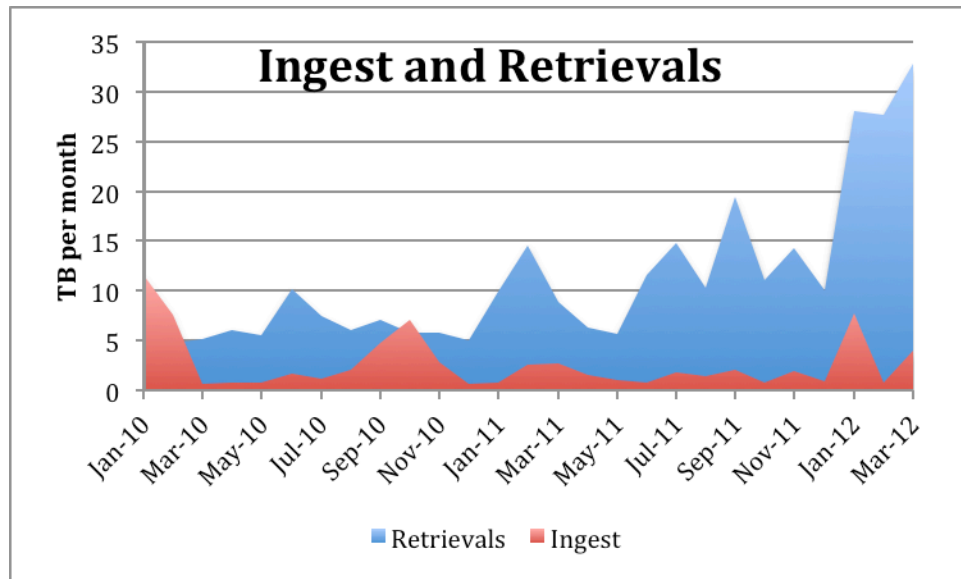
Overall Mission: MAST supports active and legacy mission datasets and related catalogs and surveys, focusing primarily on data in the ultraviolet, optical, and near-IR spectral regions. Support includes providing data curation, providing expert support to users of the data, providing access to data-specific 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 April 1, 2012 MAST holdings are nearly 190 TB, including over 70 TB of Hubble Legacy Archive data, 63 TB of HST standard pipeline products and over 1.5 TB of Community Contributed High Level Science Products. The figure below shows MAST holdings as of April 1, 2012.



The archives at STScI have distributed over 32 TB of data between March 2011 and March 2012, and have ingested 4 TB of data. The figure below shows the statistics on data ingest and distribution to the public from Jan 2010 through March 2011. The large spikes of ingest in January 2010 and January 2012 are 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 fifth data release (DR5) in March of 2011. In this release, the HLA offers a new interface to the footprint data, which outline the region of the sky covered by any HST dataset. With this new interface, users can easily zoom and pan in the footprint view, select multi-instrument datasets of interest, retrieve selected data and utilize complete cart and table integration. An additional improvement is clickable sources on the interactive display that show users information about the source such as magnitude. The cart has a new feature where a user can choose to download data in zip format or sequentially. A new table filtering option has been added where in the RA or Dec columns, the values may be entered either in decimal degrees or in "sexagesimal" format (hh:mm:ss, dd:mm:ss). Also, multiband source lists are available for ACS and WFPC2.

New data available at release include the completion of WFPC2 processing, including those data that were still proprietary at the time of the last release, and the addition of

multi-wavelength source lists for ACS and WFPC2 data. New sets of High-Level Science Products, including the Early-Release Science WFC3 observations and the first images from the Multi-Cycle Treasury program CANDELS, have also become available. Also, ECF-generated ACS and NICMOS extracted GRISM data are now hosted at STScI

Following DR5, work started for DR6 (released January 2012). Interface changes in this release include interactive display, search and data selection enhancements. Also, The HLA Simple Image Access Protocol server is now fully compliant with the Virtual Observatory standards. New data became available with DR6. An improved data processing pipeline produced new WFC3 image data (62% of data from observations that have a release date up to 11-30-2011). New sets of High-Level Science Products were released including products from the large Multi-Cycle Treasury Programs: CLASH (Cluster lensing And Supernova survey with Hubble), PHAT (The Panchromatic Hubble Andromeda Treasury), and CANDELS (Cosmic Assembly Near-IR Deep Extragalactic Legacy Survey). Also, additional ACS, WFPC2 and NICMOS data were released.

Galaxy Explorer (GALEX)

NASA-administered operations of the GALEX satellite were terminated in mid-February 2012. Plans were initiated to transfer the satellite to CalTech. In anticipation of this shutdown, MAST worked with the GALEX project on plans to transfer all scientific and engineering data for archiving and deep storage (respectively).

Responding to the perceived need for observers to obtain light curves and images of GALEX data over arbitrary intervals, MAST developed a plan to work with former member of the GALEX staff to convert intermediate data files to photon list files using software adapted from the Project's pipeline processing system. This conversion is computation-intensive and requires several computers at STScI to complete the project. During 2012, the former GALEX team member will create tools to extract fluxes to build images time slices and light curves from the completed photon list files..

Data Deliveries:

The GALEX project delivered GR6 grism data and addendum-imaging products (and errors in the original delivery corrected). When made public, these data completed the GR6 ingest process from 2010. The GR6 products also contained all "raw-6" files for observations through the Feb., 2012. These files are needed for conversion to the photon list.

In early 2012 most GR7 imaging products (excluding the All Sky Survey) were delivered. This included several areas of the sky that had never been observed before such as the Galactic cap, low Galactic latitudes, and Magellanic Cloud areas.

MAST received nine GI deliveries during the report period, the total size of which was 337 GB. This completed the delivery of the GI products. (Final reprocessing of the GI data were delivered to MAST under its "GII survey" program in the GR6 and GR7.

Catalogs:

MAST worked with the GCAT performing essential quality control steps to the GALEX self-matching of GALEX Catalog (GCAT) objects for the All Sky and Medium Survey Catalogs.

CasJobs:

The MAST CasJobs facility, which previously included cross matched targets from the GALEX and Kepler missions, was expanded to include infrared magnitudes from ground based surveys (2MASS and UK Infrared Telescope) of the Kepler field. In February the UBV and Kepler Isaac Newton Telescope Survey (KIS) were ingested. These surveys include new colors (and some new targets) in the Kepler field.

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.

MAST staff

MAST staff added a new field and search option to the Kepler Target Search interfaces allowing users to find targets in various categories such as "released planetary candidate", "Exoplanet Host Star", "Possible artifact", etc. Entries are listed as semi-colon separated strings allowing multiple flags per target.

MAST staff ingested and cross matched catalogs from ground-based observing groups to provide new filter magnitudes for objects in the Kepler field of view. An enhanced search page provides access to this new information.

Community interaction

Survey

In September 2011 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/2011/index.html>).

MAST Users Group

The MAST Users Group (MUG) met in September 2011. The MUG provides an essential user perspective on archive operations and development. All the presentations and the MUG report have been posted on the MAST website (<http://archive.stsci.edu/mug/index.html>).

Outreach

MAST and the HST Office of Public Outreach (OPO) continued 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 Virtual Observatory (VO) web services. OPO delivered final versions of selected images and meta-data. All required services and RSS feeds were completed and released in late 2011. The data are also available through the Virtual Observatory Discovery Portal.

The European Space Agency (ESA), OPO and the Hubble Legacy Archive (HLA) are collaborating on an outreach project called "Hubble's Hidden Treasures". Members of the public use the HLA to search for data that have not been published outside the scientific literature and construct color images to be submitted to the project's Flickr page. On the first full day after the announcement the HLA and MAST server loads were able to successfully handle load 10 to 20 times the normal level.

Virtual Astronomical Observatory (VAO)

The VAO registry-publishing interface was reworked to be more complete and usable. The database behind the registry had a major rework to accommodate changing requirements and to improve efficiency. A new VAO discovery portal was developed for the VAO. The VAO and MAST discovery portals use the same technology stack contributing to development efficiency.

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.

Spectral Classes: A MAST Tool for Selecting Stars by Spectral Type

MAST has constructed a tool that associates "best" spectral types in the literature with targets observed in MAST's spectroscopic missions. Users can access this tool by clicking "*Spectral Classifications*" in the Tools tab at the top of the MAST home page or

by navigating directly to http://archive.stsci.edu/spec_class. The tool is an interface form called "Spectral Classes of Like Stars."

Hardware and Migration

At the end of 2011 MAST purchased a significant amount of SAN storage space – almost 300TB. Of this 100TB was to replace storage hardware that was end-of-life, 100TB was to accommodate planned increases in data holdings, and 100TB was to satisfy an approved overguide request to save the GALEX pixel data that would otherwise have been lost following the closeout of mission operations.

In addition, we also purchased 4 new servers to begin replacing existing systems that are also approaching end-of-life, as well as 1 new system to accommodate expected increases in usage. The new systems are currently being configured.

Community-Contributed High-Level Science Products (HLSP)

Eleven sets of community contributed reduced science ready data sets were ingested into MAST over the past year. Two additional sets are in various stages of preparation and ingest. Three of the sets are for Hubble Multi-Cycle Treasury Programs for which there have been multiple deliveries.

The High-Level Science Products continue to be extremely popular.

