

## NASA Data Center Annual Program Plan

NASA Grant Number NNX09AF08G / STScI Grant J1346

March 2009 through February 2010

The Multi-Mission Archive at Space Telescope Science Institute (MAST) Phase II  
(Optical/UV Science Archive Research Center)

Space Telescope Science Institute

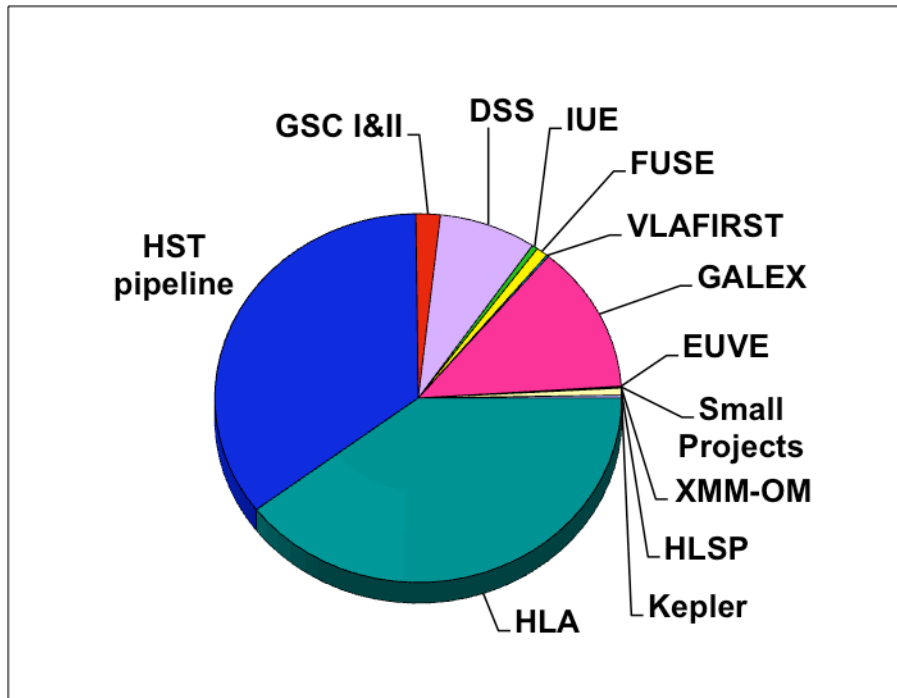
3700 San Martin Drive

Baltimore, MD 21218

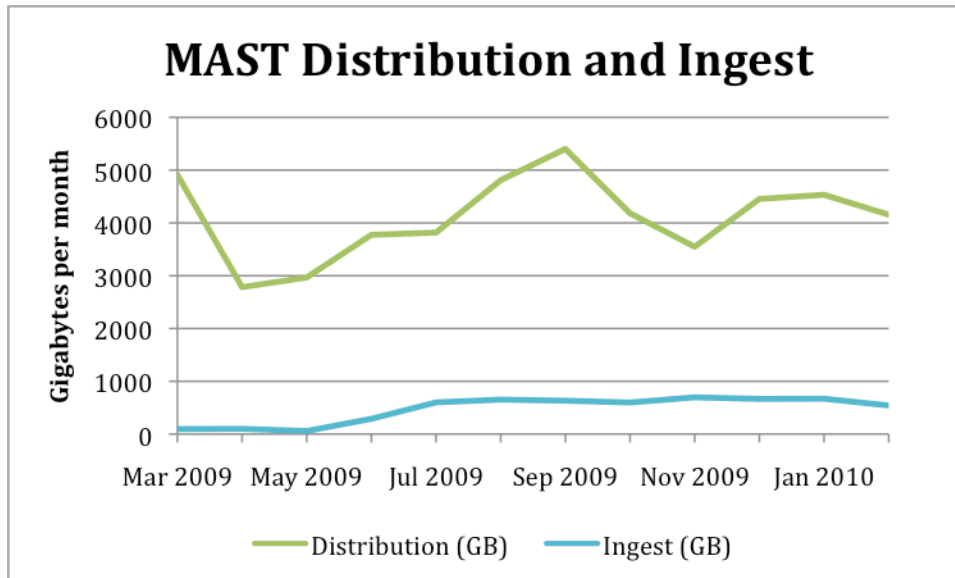
**Overall Mission:** MAST supports active and legacy mission datasets and related catalogs and surveys, focusing primarily on data in the ultraviolet, optical, and near-IR spectral regions. Support includes providing data curation, providing expert support to users of the data, providing access to data-specific 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, 2010 MAST holdings are over 125 TB, including over 42.2 TB of Hubble Legacy Archive data, 43 TB of HST standard pipeline products and nearly 460 GB 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 March 2009 through February 2010.



## Mission/Project reports

### *Hubble Legacy Archive (HLA)*

HLA had its third data release (DR3) in May 2009. In this release, WFPC2 source lists, ACS mosaic images, the first set of NICMOS drizzled images and the first set of ACS extracted grism spectra (contributed by the ST-ECF) were added to the HLA data holdings. This brought the total data holdings to approximately 34 TB. DR3 also included numerous user interface enhancements: improved cart handling capabilities with better interaction with ST-DADS, better filtering and sorting capabilities and improvements in the interactive display in the form of GALEX catalog overlays and advanced contrast controls.

Following DR3, work started on DR4, which is planned to be installed in Spring 2010. This release is targeted to contain the complete set of NICMOS drizzled images; to bring ACS products up to date; to complete the set of WFPC2 drizzled products from the final calibrations along with updating the associated WFPC2 source lists; to add new High Level Science Products (HLSP) to the HLA databases and interfaces; to add new ACS mosaics and an initial set WFC3/COS data. User interface enhancements include a new 1D-plotting tool that is

integrated into the interactive display, better source list distribution capabilities, enhanced HST based lookup capabilities, and numerous minor changes and performance improvements.

### *Galaxy Explorer (GALEX)*

The GR4 data release was supplemented with GR5 adding depth (or exposure) and sky area for the deeper surveys; the all sky survey was completed with the GR4 in 2008. A number of GR4 Grism sky areas ("tiles") did not have correct imaging antecedent files, and MAST worked with Caltech to identify these tiles so that it could reprocess them correctly. MAST received a monthly delivery of Guest Investigator data products, ingested these products and provided the GI office at Goddard a detailed listing of the new products. (The GI office notifies the PIs of the availability of their proprietary data as an outcome of this process.)

MAST has begun preparations for delivery and ingest of GR6. Several changes to schema are required so that ingest of GR6 can proceed automatically. In addition, during the year MAST participated in weekly discussions with the project regarding the GALEX catalog of far-near UV object cross-matches for the all sky and medium deep surveys, as well as of cross matches between these surveys called "GCATT". MAST has provided advise to the GCATT group regarding documentation of this new catalog.

MAST and JHU's Sloan Digitized Sky Survey group collaborated to produce GALEX-Sloan object cross-match tables, using the GR4/5 catalogs. The final cross-match tables were delivered from JHU and installed as a new search application in the CasJobs tool. MAST wrote a tutorial describing use of the CasJobs cross-match. An upgraded version of JHU CasJobs interface was imported and configured to utilize the GR4/GR5 dataset.

### *Far Ultraviolet Spectroscopic Explorer (FUSE)*

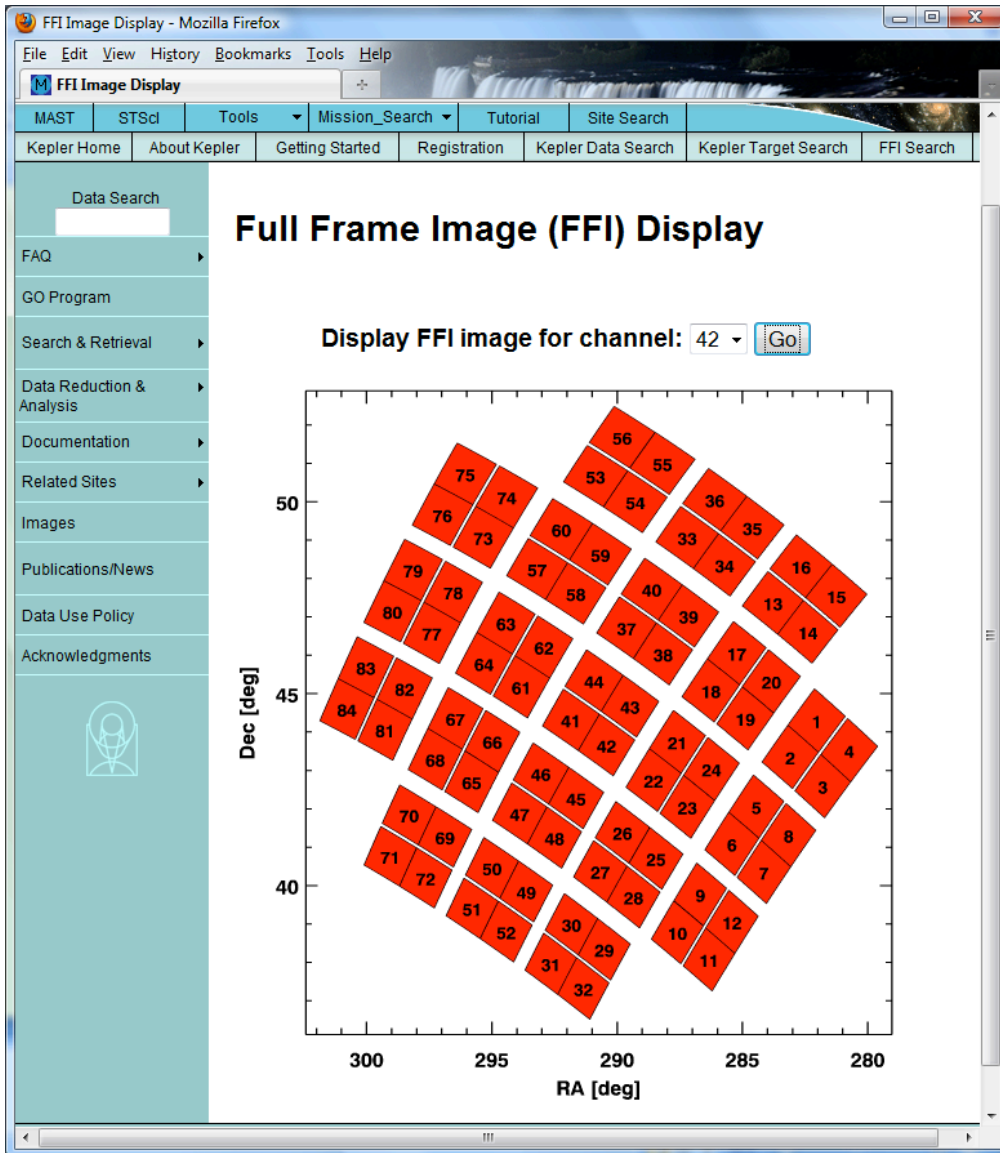
The FUSE project completed final processing the entire FUSE archive, and these standard products were archived in the archive. MAST staff members worked closely with the FUSE team to identify, archive, and integrate additional products and documentation that were available on the FUSE website into the MAST/FUSE website and archive. This work was generally completed in May 2009, although the reprocessing effort was completed in the fall of 2009.

### *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 collaborated to design and implement several search interfaces and website documentation and content. Kepler project funded the work needed to implement Kepler interfaces and the website using existing MAST web technology and interfaces, just as the HST archive ingest and cataloging system was leverage to produce the Kepler archive pipelines.

MAST staff utilized astrometry.net to determine World Coordinate Systems (WCS) for the Kepler Full-Frame Images (FFI) that do not yet have WCS assigned as part of the processing pipeline. After the WCS were assigned, MAST staff utilized the FITS display software developed for the HLA to display the FFIs and to overlay catalog information on the image. Below is a screen shot of that interface.



## *Epoxi/EPOCH*

MAST participated in the Planetary Data Systems EPOXI mission review. This meeting was held to audit the quality and contents of mission data products for this supplementary program of the Deep Impact spacecraft. These products consist of images and spectra of the Earth seen at a distance and images of stars that host exoplanets. MAST has prepared a mission site and the supporting documentation. By the end of the reporting period the site was nearly ready, and MAST now awaits light curves for two of the exoplanet stars, as well as a release date embargo of February 28, 2010.

## *CHIPS*

Representatives of the CHIPS project delivered FITS data products for this small mission, and MAST readied supporting documentation. A prototype version of this website was prepared.

## **Community interaction**

### *Survey*

In May 2009 MAST administered what has become a yearly survey to gather feedback about our service and to gauge priorities for future work. There were 299 respondents to the survey. The results and many of the comments have been posted on the MAST website. (<http://archive.stsci.edu/surveyresults/2009/index.html>)

### *MAST Users Group*

The MAST Users Group (MUG) met in July 2009. 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>)

### *Newsletter*

MAST distributed two electronic newsletters during the past year. All MAST newsletters are available online. ([http://archive.stsci.edu/archive\\_news/](http://archive.stsci.edu/archive_news/)).

## **Outreach**

MAST, in coordination with the HST Office of Public Outreach (OPO), has begun defining the work needed to integrate HST Press Release images and meta-data into the archive. Plans include building a database and completing any required meta-data (e.g. coordinates). These data will all be made available through Virtual Observatory (VO) services.

## **Virtual Observatory (VO)**

A new version of the Virtual Observatory registry was deployed in June 2009. Staff members collaborated with other VO registry providers to provide curation of the registry contents and content harvesting activities, throughout the year. One representative attended the November 2009 International Virtual Observatory Interoperability meeting. Staff members participated in review and requirements definition for the Table Access Protocol (TAP) and the TAP Data Model.

### **Other Major work efforts**

MAST staff worked on many projects during the past 10 months that introduced new or enhanced capabilities or attributes. We describe a few highlights below.

#### *Galexview GalexView / Galex Map Browser*

The flash interface galexView v. 1.4 was deployed with improved performance, a screen shot tool, and improved color palette to denote object positions. In addition to searching sky tiles from the mission surveys, galexView is now also able to access the Gil de Paz GALEX atlas of nearby (~1000) galaxies.

A library of Flash-based graphical components was developed including a Name Resolver, a Results Grid/Spreadsheet and an Image Viewer. These components are now utilized by galexView and were reused for an HST interface that replaces the old Starview Java application. The Archive Component Library will continue to be enhanced and extended to other missions in 2010.

Following a decision to integrate the GALEX Map browser and galexView search tools, MAST used the Google Maps API to create several sets of GALEX jpeg images representing a mosaic across the sky, each of a different resolution and for each of the GALEX surveys. A new version of galexView flash interface will seamlessly access these sky-wide maps. For a given sky area the deepest exposure will be rendered. The initial version the images are rendered in Mercator projection so the sky cannot be observed at a declination within 5 degrees of the poles. A later version will address this limitation by using a TOAST projection.

#### *Hardware and Migration*

Over the last year MAST purchased a number a new servers and disk storage, both to replace some aging hardware and to expand data holdings whilst also improving reliability and availability of services. We used to have different servers for various MAST datasets, and since some were older systems we were having more downtime due to hardware issues. We are now transitioning our data and services to modern clustered systems in a failover configuration so that any failure will not affect the data availability and services we provide to users. In addition, approximately 80TB of disk were purchased for placement in a new SAN also to improve reliability and availability. About 50% of this was to replace old, failing disk and the remainder was to augment storage for both GALEX data releases and HLA products. Work is well underway to migrate data holdings to the new storage and to migrate the databases to the servers.

During this reporting period, many of the MAST database tables were migrated from the Sybase database to a MS SQLServer database. The initial work was instigated when

large Kepler and HST searches cause the server to slow and sometimes to hang. SQLServer accommodates several indexing options that improve search performance, which MAST DBAs had used for HLA and GALEX searches. Using the SQLServer improved performance by as much as a factor of 100.

#### *Spectra in the Virtual Observatory (VO)*

MAST staff members completed additional “spectral container” fits data files and associated web services to meet the criteria of the VO Simple Spectral Access Protocol (SSAP), v1.1. All SSAP services were registered in the new VO registry.

#### *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. Five additional sets are in various stages of preparation and ingest. 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. Below are some sample product images, including the Carina Nebula and the Andromeda Galaxy.

