*Astronomical Data Analysis Software and Systems XVIII* **O2.3**
*ASP Conference Series, Vol. XXX, 2009*
*J. Lewis, R. Argyle, P. Bunclarck, D. Evans, and E. Gonzales-Solares, eds.*

# Design of the Spitzer Space Telescope Heritage Archive

Deborah Levine[1], X. Wu, J. Good, A. Alexov, B. Berriman, P. Capak, S. Groom, T. Handley, J. D. Kirkpatrick, M. Lacy, M. Legassie, S. Laine, L. Ly, T. Roby, Russ R. Laher

*Infrared Processing and Analysis Center, California Institute of Technology, M/S 100-22, Pasadena, CA 91125*

**Abstract.** It is predicted that Spitzer Space Telescopes cryogen will run out in April 2009, and the final reprocessing for the cryogenic mission is scheduled to end in April 2011, at which time the Spitzer archive will be transferred to the NASA/IPAC Infrared Science Archive (IRSA) for long-term curation. The Spitzer Science Center (SSC) and IRSA are collaborating to design and deploy the Spitzer Heritage Archive (SHA), which will supersede the current Spitzer archive. It will initially contain the raw and final reprocessed cryogenic science products, and will eventually incorporate the final products from the Warm mission. The SHA will be accompanied by tools deemed necessary to extract the full science content of the archive and by comprehensive documentation.

## 1. Introduction

This paper will describe the software design of the archive, which integrates the archive holdings into the current IRSA infrastructure, and, at the same time, offers powerful tools that allow users to explore the full range of scientific content of the Spitzer data sets.

## 2. About Spitzer Space Telescope

NASA's Spitzer Space Telescope is an infrared (IR) observatory built to study the creation of the universe, the formation and evolution of galaxies, the formation of stars and planets, and the chemical evolution of the universe. Spitzer was launched on 2003 August 25.23 UT into an Earth-trailing solar orbit.

Spitzer's 85-cm cryogenically cooled beryllium Ritchey-Chretien telescope operates at temperatures ranging from 5.5 to 10 K. Its three cryogenically cooled Science Instruments perform imaging and spectroscopy in the 3-180 $\mu$m wavelength range: Infrared Array Camera (IRAC), Infrared Spectrograph (IRS) and Multiband Imaging Photometer for Spitzer (MIPS). Spitzer's goal lifetime of 5 years has already been exceeded, and it is now predicted that the cryogen will be expended in April 2009. The two shortest wavelength arrays of the IRAC instrument can take quality data after the cryogen is spent and the observatory

[1]E-mail: deblev@ipac.caltech.edu

warms up, assuming 2 years of post-cryogenic Spitzer Space Telescope operations have been funded.

Only one instrument is operable at a time and, depending on which instrument is observing and how the observations have been structured, the amount of raw science data generated daily can range from around 0.3 to 2 GB per day. The archive will contain about 25 TB of reduced data at the end of the cryogenic mission. The cryogenic archive will be handed over to IRSA for long-term curation in April 2011.

## 3. Why a Heritage Archive

The current Spitzer archive is served through a Java application called Leopard, which accesses the current archive database, which is a copy of the science operations database that is replicated during SSC image processing. This database was built to support uplink and pipeline operations and operations security, rather than with an emphasis on archival research, very long term maintenance, and ease of migration. In addition, Leopard was optimized for the identification and retrieval of data from specific observations by the observer or researcher familiar with the observation. The SHA will have a simplified database schema that is more optimal for archive usage, additional functionality, and an improved interface to facilitate deeper and more complex mining of the archival data.

The SHA will be built by SSC and maintained long-term by IRSA. Close collaboration between SSC and IRSA staff is needed, and a SHA IPT (Integrated Product Team) was formed and regular meetings established. Working groups containing members of both organizations were formed and requirements and design reviews were held with oversight boards including members of both organizations user panels.

## 4. Holdings

The SHA will contain ≈25 TB of cryogenic data when it is handed over to IRSA. The SHA holdings will include:

1. Spitzer raw data
2. Spitzer pipeline data:
    - Science quality data products
    - Ancillary data
3. Calibration products
4. Data quality analysis information
5. Observation design information
    - Program title & abstract
    - Observational constraints
6. Enhanced Spitzer data and ancillary data
    - Data from the legacy teams
    - Includes catalogs, models, etc.
7. SSC-produced "enhanced" data

The SHA will also eventually contain the IRAC data from the "warm" mission, which will be added incrementally.

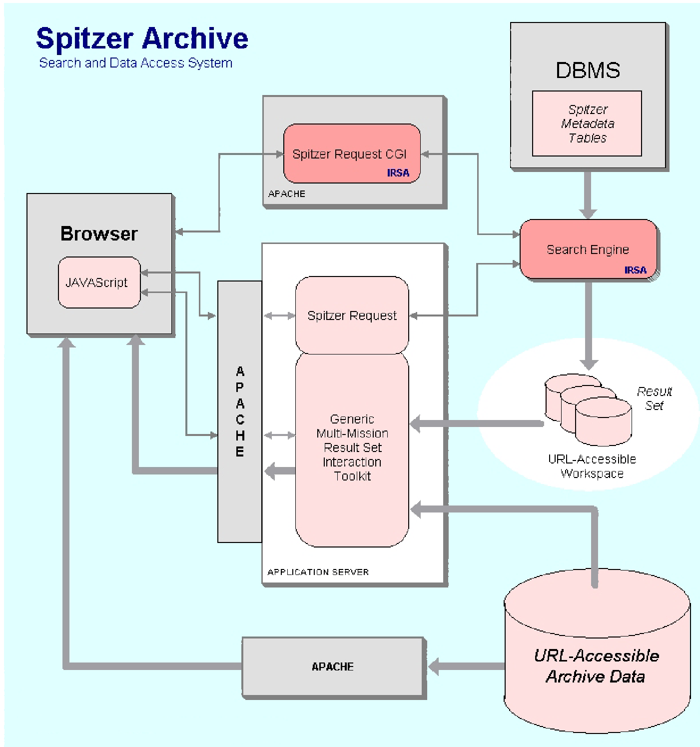


Figure 1. Overview of the Spitzer Heritage Archive System.

## 5. Design Drivers

The requirements for the SHA were derived from science use cases, which were collected from a variety of active Spitzer users and instrument experts. In addition, it was important to design the SHA as an SSC/IRSA collaboration; both organizations will have responsibility for the final system and the data in it, and we will have leveraged software and expertise in both organizations.

The resulting design attempts to address the full range of user expectations and limitations by providing a powerful tool to mine the Spitzer data while also permitting access to basic functions through simple interfaces.

An emphasis was also made on laying the groundwork for future evolution by designing and developing maintainable and well documented software that is sufficiently modular and scalable to facilitate the ability to swap in new technologies and adapt it to additional datasets.

Finally, the SHA must be built without interfering with ongoing operations.

## 6. Architecture and Key Technologies

Figure 1 provides an overview of the SHA system.

The design incorporates independent components with clearly defined interfaces that are compatible with IRSA and VO (virtual observatory) service infrastructure. It uses IRSA standard database search tools and provides query logging and NASA-required data tagging, as well as supporting VO protocols.

The database has been restructured to make it relatively flat, and to eliminate stored procedures as part of an effort to ensure database independence. The data-search capability is relatively isolated to support upgrades and VO protocols are used for data access external to the SHA system itself.

AJAX (Asynchronous JavaScript and XML) web technologies will be used to provide the user with more functionality to manipulate result sets and related data within the archive interface. Historically, archive interfaces have supported basic data location and sub-setting, but have been weaker at more involved tasks such as drill-down, data selection, and interactive visualization. Using modern web tools and techniques, it is now feasible to make a web browser interface much more powerful. In particular, the SHA web site will implement advanced table manipulation and FITS visualization. Its web-based FITS viewer will be an AJAX-based tool with functionality similar to installed applications, such as interactive pixel value/coordinate readout and manipulation of display parameters. The display will be melded seamlessly with fully scalable server-side database technology.

Specialized tools are also being developed on the server end to provide extremely fast spatial searches for both catalog and image coverage data. Based on R-Tree indexing, these tools were originally developed for the National Virtual Observatory and have been used there for pan-archive inventory searches.

The above toolkits will be both general and portable. IRSA plans to use the table-interaction toolkit with all its tabular result sets, and the spatial indexing software will be part of an upcoming IRSA release.