

Wide-field Infrared Survey Explorer (WISE)

WSDC Archive Design

Version 1.0

6-October-2008

Prepared by: Lin Yan



**Infrared Processing and Analysis Center
California Institute of Technology**

WSDC D-D007

Concurred By:

Lin Yan, WISE Science Data Center Lead Archive Scientist

Roc Cutri, WISE Science Data Center Manager

Revision History

Date	Version	Author	Description
2/26/08	0.1	R.M. Cutri	Initial Draft
11/06/08	1.0	Lin Yan	Version 1

Table of Contents

1	INTRODUCTION	7
	1.1 Driving Requirements.....	5
2	OVERVIEW.....	7
	2.1 Scope of this document	7
	2.2 Applicable Documents.....	7
3	FUNCTIONALITY OF WISE/IRSA DATA ARCHIVE.....	8
4	WISE DATA PRODUCTS SERVED BY IRSA.....	9
	4.1 Data Products for general users.....	10
	4.2 Data Products for the WISE Science Team.....	12
5	IMPLEMENTATION.....	16
	5.1	
	5.2	
	5.3	
6	STAFFING PLAN.....	17

1 INTRODUCTION

The Wide-field Infrared Survey Explorer (WISE) will carry out an all sky survey in 3.3, 4.7, 12 and 23 micron mid-infrared bands. WISE will produce a digital Image Atlas and a Source Catalog that will contain approximately 300 million objects. These products will enable a broad variety of research ranging from studies of solar system asteroids, the closest stars and brown dwarfs to the most luminous, and distant galaxies in the Universe. The WISE science data will serve as an important reference for planning observations and interpreting data obtained with future ground and space-borne observatories such as TMT and JWST.

WISE is designed to achieve a minimum point source sensitivity on the ecliptic corresponding to flux signal-to-noise ratios greater than 5 at flux densities of 0.12, 0.16, 0.65 and 2.60 mJy at 3.3, 4.7, 12 and 23 microns, respectively, in regions of the sky not confused by Milky Way stars and diffuse emission. WISE is scheduled to launch in November 2009 and will have an in-orbit checkout (IOC) phase of one month, followed by a six month baseline on-orbit data acquisition operations period. A preliminary Source Catalog and Image Atlas constructed from data acquired from the first 50% of the sky surveyed will be released six months after the end of the on-orbit operation phase. The final Catalog and Atlas will be released 17 months after the end of on-orbit operations.

The Infrared Processing and Analysis Center (IPAC) at the California Institute of Technology is the WISE Science Data Center (WSDC). The WSDC is responsible for ingesting raw wise science and engineering data and ancillary files from the WISE Mission Operations system, processing those data into the calibrated image products, archiving those products and distributing them to the Scientific and Public Communities.

This document describes the plan and implementation of the WSDC ARCHIVE subsystem that is responsible for the storage and distribution of the WISE image and tabular data during and following the WISE mission.

1.1 Driving Requirements

Table 1 contains a listing of the requirements derived from the WSDC Functional Requirements Document (WSDC D-R001) that define the design and functionality of the WISE Archive.

Table 1 - WISE Archive Driving Requirements

ID	Requirement
L4WSDC-050	The WSDC shall create a copy of the Level 0 science data in a medium appropriate for permanent long-term storage.
L4WSDC-051	The WSDC shall make the WISE catalog and image products available to the community via the internet through appropriate web-based tools.
L4WSDC-052	As a goal, the WSDC will maintain the data products in a way that distribution of the complete WISE Source Catalog to users via portable

	media would be possible.
L4WSDC-053	The WSDC shall make the Image Atlas and Catalog products accessible to the astronomical community in collaboration with the NASA/IPAC Infrared Science Archive (IRSA) to ensure long-term availability beyond the end WISE missions operations and data processing phase, and to insure interoperability with other NASA mission archives.
L4WSDC-054	The WSDC shall maintain a complete copy of the WISE science data set and software source code at a secure off-site location during the WISE mission to ensure survivability in case of major catastrophe.
L4WSDC-055	After the WISE mission, a copy of the Level 0 science data shall be delivered to the National Space Science Data Center (NSSDC) for permanent archive.
L4WSDC-056	The WSDC shall maintain an archive of metadata derived from data processing for the individual science images for the duration of the project for the purpose of analysis and support of image access tools.
L4WSDC-057	The WSDC shall provide an online repository for operations products for the life of the project.
L4WSDC-058	The MOS shall capture and archive the following data sets received or created by the EOS during the mission: (a) all telemetry data received on the ground, (b) all commands sent to the spacecraft, (c) all sequence products, (d) all data processing logs.
L4WSDC-059	Sample WISE images shall be made available for outreach purposes within 1 month of start of normal operations.
	2.2.3.1 Data Access
L4WSDC-060	The WSDC archive shall provide a web-based interface to enable selection, display and retrieval of any or all single-epoch images and combined Atlas Images based on position or time of observation for the purpose of quality assurance, validation and analysis. The goal shall be to also allow image selection on any image metadata parameter.
L4WSDC-061	The WSDC archive shall provide a web-based interface to enable selection of sources extracted from single-epoch frames and/or combined Atlas Images based on position, flux, or combinations of any parameter maintained in the extracted source databases or Source Catalog.
L4WSDC-086	The web-based interface to the WISE Image Atlas shall allow the user to view and retrieve an image in any of the four WISE bands with any specified center (tangent point) and any size up to at least 1°x1°.

2 OVERVIEW of the WISE ARCHIVE

WISE science data products will be distributed to the community via the on-line and computer-compatible services of the NASA/IPAC Infrared Science Archive (IRSA) at IPAC. To efficiently distribute the science data, the WISE Archive Subsystem will provide online tools and Application Programming Interfaces (APIs) that enable flexible and easy access to various data products.

The WSDC Archive Subsystem has three major components:

- **IRSA archive** – All processed data, derived metadata and the final science data product will be stored, maintained and distributed to users by IRSA. All public data products are located at IRSA. Existing or slightly modified IRSA access tools and Application Programming Interfaces (APIs) can be used to serve these data. IRSA will serve the Project Science Team and the general public via on-line interfaces, and will interoperate with other NASA and general archive services via machine-friendly interfaces.
- **Operations archive** – The operations archive will hold raw data, processed data and the derived metadata within the processing system. The storage is physically located close to the processing cpu's to minimize latency due to network I/O. The operation archive serves data to the processing pipelines, QA system and Final Product Generator (FPG) operations through direct file access. Databases are distributed in a predetermined directory structure on the WSDC data servers. Access to the Operations archive will be limited to processes originating within and users physically located at the WSDC.
- **Miscellaneous archives** – These include the engineering database, ground test data, and the Mission Operation System (MOS) archive.

2.1 Scope of this Document

The WISE archive has three major components, and this document focuses on the WISE data archive at IRSA. This document describes the functional requirements for the WISE Science Data Center Archive Design. We will discuss in details what types of data the WISE archive will contain and what functionalities it will have to serve the general astronomical community and the WISE project team.

2.2 Applicable Documents

The WISE operations archive and other archives are also discussed in the following documents.

- WISE Science Data Center Function Requirements Document (FRD - WSDC D-R001)
- WISE Science Data System Functional Design Document (FDD - WSDC D-M004)
- WISE Project Data Management Plan (PDMP - WSDC D-M002)

- WISE Science Data Quality Assurance Plan (WSDC D-M004)

3 FUNCTIONALITY OF WISE/IRSA DATA ARCHIVE

In this section, we discuss the functional capabilities the WISE/IRSA data archive shall offer to the general users and the WISE project team.

- ◆ **Data Query and Delivery via IRSA:** The WISE/IRSA shall serve the WISE catalog and image products to the community and the project science team via the internet through appropriate web-based and computer-friendly interfaces.
 1. **Catalog Search:** The Catalog interface shall allow searches of sources extracted from single-epoch images and/or combined Atlas images based on position, flux, flux ratio, or combinations of any parameter maintained in the extracted source databases or source catalogs. **The Catalog interface shall also enable queries on any metadata table maintained in the Archive.** The interface shall allow complex and efficient queries of massive catalogs and metadata tables.
 2. **Image Services:** The WISE/IRSA archive shall have a web-based interface and computer-friendly interfaces that provide image services to general users. This interface shall enable selection, display and retrieval of any or all single-epoch images and combined Atlas Images based on position and time (in the case of the single-epoch images) of observation using the information stored in the image metadata. The returned images will be in the pixel format and projection of the stored images. This image service shall allow both interactive and batch requests.
 3. **Image Pixel Server:** The web-based interface shall allow the user to view and retrieve an Atlas Image in any of the four WISE bands with any specified center and any size up to at least 2 by 2 sq.degrees (TBD). This includes *image cutout and FinderChart modes*. The cutout tool returns small “cutouts” of a user specified size of image data sets. The returned data can be in FITS and JPEG formats. The FinderChart tool is a visualization tool to cross-compare WISE image with other datasets such as 2MASS and SDSS. It displays a finder chart showing the same region of the sky at various wavelengths and epochs, and supports user interaction with the images.
 4. **Image Inventory Service:** This service shall provide a listing of all WISE images and their metadata that cover a user-specified point or region on the sky.

- ◆ **The Long Term Availability:** The WISE/IRSA archive shall ensure the long-term availability of the preliminary and the final science Image Atlas and Source Catalog products to the general astronomical community, beyond the end of WISE mission operations and data processing phase. At the completion of WISE project activities, IRSA will assume curation of the data products and explanatory documentation.
- ◆ **Interoperability with other NASA mission data and archives:** Because the WISE data products are integrated into the IRSA infrastructure, WISE data will automatically be visible to a number of IRSA data inventory and cross-mission data discovery services, including OASIS and RADAR. IRSA shall also make computer-friendly interface access available to WISE products by other NASA archives.
- ◆ **Data Delivery via Portal Media:** As a goal, the WISE archive shall be able to distribute the complete WISE Source Catalog to users via portable media such as tape.

4 WISE DATA PRODUCTS SERVED BY IRSA

The WISE/IRSA data archive will serve the data to the general astronomical community as well as the WISE project science team. The archive is a multi-tier system that stores the processed images, source catalogs and metadata. Below are the data products to be made available at the WISE/IRSA archive. Table 2 summarizes all of the planned WISE/IRSA images and tabular holdings. Section 4.1 and 4.2 discuss in detail each of the data products.

Table 2 – Summary of WISE/IRSA image and tabular holdings

Product name	Description	Access	No. Entries
IMAGES			
Preliminary Image Atlas (science, coverage and noise images)	Calibrated images in the four WISE bands covering the first 50% of the sky surveyed.	General users	29,000
Final Image Atlas (sci, cov, noise)	Calibrated images in the four WISE bands covering the entire sky.	General users	57,000
L0 Images (sci, noise, mask)	These images are raw images, including science, mask and noise images for four bands.	Science team only	18 millions
L1b Images (sci, noise, mask)	These are single epoch, calibrated image, including science, mask and noise images for four bands.	Science team only	18 millions
L3 Images (sci, cov, noise)	These are the final coadded, calibrated images. These include science, coverage and noise images in four bands.	Science team only	57,000
TABLES			
Preliminary Source Catalog	Information on source positions, fluxes, uncertainties and SNR derived from the preliminary images.	General users	150 millions
Final Science Catalog	Source catalog derived from the final science images	General users	300 millions

Atlas Image Metadata	Information about each Atlas image (likely a subset of L3 Image Metadata)	General users	57,000
Atlas Image Tiles Metadata	Information about each pre-defined Atlas Image tile. Center, footprint, number of images contributing, etc.	General users	57,000
L0 Image Metadata	Information about each Level 0 image frame	Science team only	18 millions
L1 Image Metadata	Information about each Level 1b image frame	Science team only	18 millions
L3 Image Metadata	Information about each Level 3 image frame	Science team only	57,000
L1b Source Working Database	Extracted source information from Level 1b images. 4-band, photometrically and astrometrically calibrated entries.	Science team only	100 millions
L3 Source Working Database	Extracted source information from combined Level 3 images. 4-band, photometrically and astrometrically calibrated entries.	Science team only	300 millions
Calibration database	Contain all calibrators, primary & secondary.	Science team only	500
Calibration table	Contain match information for all calibrators during the mission.	Science team only	10 million
Scan metadata table	Contain the final zero points for each scan	Science team only	10,000
Solar System Objects	Information about known solar system objects that fall within the footprint of L1b images at the time of WISE observation. WISE source extractions for associations.	General users	1 millions
Predicted Events File	List of significant events predicted to occur, including scan, time, etc.	Science team only	100,000
Reconstructed Events File	List of significant events that actually occurred, along with actual scans, times, etc.	Science team only	100,000

4.1 Data Products for the General Users:

The two major data deliveries from the WISE mission are the preliminary and the final science data products.

- **Preliminary Science Data Products**

The preliminary data products are derived from the first 50% of the sky surveyed by the Mission. These preliminary products will be released 6 months after the end of on-orbit data collection.

1. **Image Atlas:** Calibrated images in the four WISE bands covering the first half of the sky surveyed. Photometric and Astrometric calibration will be tied to the WISE Source Catalog. Images in all four WISE bands will be resampled to the common pixel grid and scale. The final Image Atlas is expected to contain roughly 19,000 image tiles, each with a size of $1.56^\circ \times 1.56^\circ$ and a pixel scale of $1.375''$ per pixel.

- ◇ Combination of all available exposures covering each field
- ◇ Depth-of-coverage maps for each image. These coverage maps shall show the number of independent observations that go into each pixel of the Image Atlas in each band. The coverage numbers shall take into account focal plan coverage and losses due to poor data quality, low responsivity and/or high noise masked pixels, and pixels lost due to cosmic rays and other transient events.
- ◇ Noise map for each combined image
- ◇ Metadata for the combined image

2. Source Catalog: This preliminary Source Catalog is derived from the WISE preliminary image atlas that contains sources detected in unconfused regions in the first 50% of the sky that is surveyed. This catalog will contain astrometry accurate to 0.5'' in respect to 2MASS All-sky Point Source Catalog on each axis for sources with SNR>20 in at least one band.

- ◇ The Catalog shall contain sources detected in the unconfused regions of the combined images that have SNR > 20 in at least one WISE band.
- ◇ The Catalog will have flux errors
- ◇ The Catalog will include in-band fluxes and flux-limits for objects detected in at least one band
- ◇ The Catalog will contain quality flags describing the detection and photometry measurements
- ◇ The Catalog will contain artifact flags, goodness of photometry measures, for examples, χ^2 , data ranges, and QA scores etc.

2. Atlas Image and Survey Metadata: The archive will contain various metadata tables enabling efficient searches of images and sources of the WISE database by users. These metadata tables shall **include Quality Assurance scores** associated with the data.

3. Explanatory Supplement: The preliminary release Image Atlas and Source Catalog will be accompanied by a preliminary Explanatory Supplement. This document provides the description of the mission, spacecraft, instruments, operations, data quality processing and characteristics of artifacts to allow their scientific exploitation.

- ◇ Documentation describing the WISE mission, spacecraft, instruments and operations
- ◇ Documentation on data products delivered to the general users
- ◇ User's guide on how to use WISE/IRSA archive and data formats

- ◇ Documentation on operation processing, data quality control, and description of data artifacts

- **Final Science Data Products**

The final WISE science data products will have the same requirements as described above on the preliminary science data products. These products will be released to the general public after 17 months of the end of on-orbit operation.

1. **Image Atlas:** The WISE/IRSA archive shall serve the general public users a digital Image Atlas that combines multiple survey exposures at each sky position. Photometric and Astrometric calibration will be tied to the WISE source catalog. Images in all four WISE band images will be resampled to the common pixel grid and scale.

- ◇ The final Image Atlas is in FITS format. It is expected to contain roughly 19,000 image tiles, each with a size of $1.56^{\circ} \times 1.56^{\circ}$ and a pixel size of $1.375''$.
- ◇ Combination of all available exposures covering each field
- ◇ Depth-of-coverage maps for each image. These coverage maps shall show the number of independent observations that go into each pixel of the Image Atlas in each band. The coverage numbers shall take into account focal plane coverage and losses due to poor data quality, low responsivity and/or high noise masked pixels, and pixels lost due to cosmic rays and other transient events.
- ◇ Noise map for each combined image
- ◇ Metadata for the combined image

2. **Source Catalog:** This final Source Catalog contains sources detected over the entire sky. (This catalog will contain astrometric accurate to $0.5''$ in respect to 2MASS All-sky Point Source Catalog on each axis for sources with $\text{SNR} > 20$ in at least one band.

- ◇ Flux measurements shall have SNR of 5 or more for point sources in the four WISE bands.
- ◇ The Catalog will have flux errors
- ◇ The Catalog will include in-band fluxes and flux-limits for objects detected in at least one band
- ◇ The Catalog will contain quality flags describing the detection and photometry measurements
- ◇ The Catalog will contain artifact flags, goodness of photometry measures, for examples, χ^2 , data ranges, and QA scores etc.

2. **Atlas Image and Survey Metadata:** The archive will contain various metadata tables enabling efficient searches of images and sources in the WISE archive. These metadata tables shall **include Quality Assurance scores** associated with the images. The various types of QA scores are discussed in detail in the WISE WDSC Quality Assurance Plan Document (WDSC D-M004).
3. **Explanatory Supplement (L4WSDC-003):** The final Image Atlas and Source Catalog will be accompanied by a preliminary Explanatory Supplement. This document provides the description of the mission, spacecraft, instruments, operations, data quality processing and characteristics of artifacts to allow their scientific exploitation.
 - ◇ Documentation describing the WISE mission, spacecraft, instruments and operations
 - ◇ Documentation on data product delivered to the general users
 - ◇ User's guide on how to use WISE/IRSA archive and data formats
 - ◇ Documentation on operation processing, data quality control, and description of data artifacts

4.2 Data Products for the WISE Science Team:

In order for the WISE Project Science Team to quickly evaluate the observations and the data analyses, the WISE/IRSA archive is set up to have an internal, password controlled access to intermediate stage data products in addition to the final science products described in the previous section. This access interface will use metadata tables to identify the physical locations of the data products, whether they reside in the IRSA database or in the WISE operation archive database.

Here we list various data the WISE/IRSA archive shall provide to the Science Team.

1. **Metadata:** The archive will contain various metadata tables enabling efficient searches of images and sources of the WISE database by users.
 - ◇ **L0 Image Metadata:** The archive will contain Level 0 image metadata. Level 0 images are images with appropriate fits header information --- metadata.
 - ◇ **L1 Image Metadata:** The archive will also contain Level 1b image metadata which contains information about each calibrated image. Level 1b images are

instrumental, astrometric and photometrically calibrated images with the calibrated parameters stored in the fits headers.

- ◇ **L3 Image Metadata:** Level 3 images are the combined, re-sampled and distortion corrected images based on multiple calibrated images.
 - ◇ **Atlas Image Metadata:** This metadata will contain information about the atlas images, likely to be a subset of L3 image metadata.
 - ◇ **Scan Metadata:** The table has the priori and derived information about each survey scan
 - ◇ **Image QA Scores:** Level 1b and Level 3 image metadata will also include image QA scores. For the WISE Project Science Team, more QA scores will be visible than what are shown to the general users in the WISE/IRSA GATOR interface.
2. **Images:** With the metadata tables, users can also access the following image products on both IRSA and WISE operation archive database.
- ◇ **Level 0 Images:** Level 0 images are raw images with cleaned-up telemetry.
 - ◇ **Level 1b Images:** These are the single epoch, calibrated images prior to co-addition (this is also called **Single Frame Products (L4WSDC-024)**).
 - ◇ **Level 3 Images:** These are the final Atlas images which shall be released to the public via the WISE/IRSA archive.
3. **Source Working Database:** The source working database shall contain **ALL** detections made by the operation pipeline. This initial detection source list will be delivered by the operation server and ingested into the WISE/IRSA archive. The working databases shall be built within the framework of IRSA, and the WISE Science Team and project staff will access these working archives using the same on-line IRSA search tools. Team access to the working archives and preliminary products will be restricted using password protection on the IRSA sites.
- ◇ **L1b Source Working database:** This database contains all sources extracted from the calibrated, single-epoch Level 1b images (Single Frame Products L4WSDC-024).
 - ◇ **L3 Source Working database:** This database contains all source detections made from the combined (Level 3) images and characterized on the Level 1b images.

4. Calibration Database: This database will contain all of the calibration relevant files, including:

- ◇ **Standards Database:** This database contains all of the extracted source information for each measurement of WISE standard stars made during the survey. This includes primary and secondary standard stars.
- ◇ **Calibration Table:** These are history files recording all of the calibration parameters derived at various times during the survey. The granularity of this table will most likely be at the scan level.
- ◇ **Scan Zero Point Table:** This table contains all of the final zero point for each scan.

6. Solar System Object Identification

- ◇ **Known Solar Objects:** The archive contain a list of known solar system objects that are predicted to fall within the footprint of WISE image frames at the time of the frame observation, and shall include the WISE source information for those objects that are positionally associated with source extractions in the WISE single-epoch image frame

8. Events Database: This database contains both the predicted and reconstructed event files.

- ◇ **Predicted Events File:** The archive will contain a list of significant events predicted to occur within a given WISE scan. This will include orbit and scan identifiers, spacecraft maneuvers and states, observing times, etc.
- ◇ **Reconstructed Events File:** The archive will contain a list of the significant events described in the Predicted Events File that actually occurred, along with their actual scans, times etc. The Reconstructed Events File may also contain events not in the Predicted Events File.

5 IMPLEMENTATION

5.1 Archive Load Module

The WISE operation pipeline will generate science data products, which then need to be moved to the IRSA database for serving to users. A specific archive load module will be developed for this purpose. The detailed functionality of this module is discussed as follows.

- **Gathering pipeline output and staging for load**

The archive load module will first prepare the pipeline output on the operation servers, including reformatting the data and adding any additional header information when necessary. Some of the data, for example, Level 1b working database, which will be served by IRSA do not necessarily need to be moved to the IRSA disk servers. Instead, these data can be stored directly on the operation pipeline servers. The archive load module will gather the location information for these data, and allow the IRSA database to directly access the WISE operation pipeline disks.

- **Database Loading**

[TBD]

5.2 Verification Plan and Maintenance

[TBD]

5.3. Development Schedule

The current schedule for the development of the WISE archive is as follows.

- **Product Specifications:**
 - Preliminary definition: Oct 3, 2007
 - Update version 1: December 3, 2008
 - Update version 2: Nov. 11, 2009
 - Update version 3: April 21, 2010
- **Schemata Definition:**
 - Preliminary: December 3, 2008
 - Update Version 1: March 18, 2009
 - Update Version 2: December 16, 2009
 - Update Version 3: May 19, 2010
- **Database Ingest Software Development:**
 - Prototype: March 18, 2009
 - Update: August 5, 2009
 - Update: December 30, 2009
 - Update: September 20, 2010
- **Interface Development:**
 - **Preliminary:** April 8, 2009
 - **Table Interface:** June 30, 2009
 - **Specification definition:** June 30, 2009

- **Update version 1:** December 3, 2009
- **Update version 2:** December 3, 2010
- **Frame Image Viewer/Access Interface:**
 - **Specification definition:** Nov 3, 2008-June 30, 2009
 - **Preliminary:** May 6, 2009
 - **Update version 1:** December 3, 2009
 - **Update version 2:** December 3, 2010
- **Atlas Image Viewer/Access Interface:**
 - **Specification definition:** Nov.3, 2008-June 30, 2009
 - **Preliminary:** May 6, 2009
 - **Update version 1:** December 3, 2009
 - **Update version 2:** December 3, 2009

6 STAFFING PLAN

[TBD]

Personnel for the WISE archive will have the following roles:

- The WISE archive will be overseen by the archive scientist. The archive scientist will produce a WSDC archive plan (this document) in preparation for the development work. The archive scientist will then direct the continued progress of the archive system from development through final handoff of the data products to IRSA.
- The WISE archive will utilize the existing IRSA database. We will develop the wise archive load module.

The time scale for this staffing is envisioned as follows:

- Archive scientist: 0.5 FTE from the present until ~6 months after the final data release.
- Archive sub-system development engineer: 0.2 FTE starting ~1 yr before launch.
- Documentation and User Support.

