

Wide-field Infrared Survey Explorer (WISE)

WSDC Archive Design

Version 2.0

4-September-2009

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
12/01/08	1.1	Lin Yan	Version 1.1
02/24/09	1.2	Lin Yan	Version 1.2
03/30/09	1.3	Lin Yan	Version 1.3
09/04/2009	2.0	Lin Yan	Version 2.0

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.....	15
	5.1 Archive Load Module.....	15
	5.2 Verification Plan and Maintenance.....	21
	5.3 Development Schedule.....	21
6	STAFFING PLAN.....	28

1 INTRODUCTION

The Wide-field Infrared Survey Explorer (WISE) will carry out an all sky survey in 3.35, 4.60, 11.56 and 22.24 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 signal-to-noise ratios greater than 5 at flux densities of 0.12, 0.16, 0.85 and 4.0 mJy at 3.4, 4.6, 11 and 22 microns, respectively, in regions of the sky not confused by Milky Way stars and diffuse emission. The expected sensitivities are 0.08, 0.095, 0.76 and 3.7 mJy at the corresponding four bands. WISE is scheduled to launch in December 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	Volume in TB
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	0.5*18,241*4*3 = 109,446 tiles (half of sky)	0.5*18,241*4*192MB=7TB
Final Image Atlas (sci, cov, noise)	Calibrated images in the four WISE bands covering the entire sky.	General users	218,892 tiles	14 TB
L0 Images (sci)	These images are raw science images for four bands.	Science team only	6.6 millions images	6.6million*4.01MB=26.5TB
L1b Images (sci, noise, mask)	These are single epoch, calibrated image, including science, mask and noise images for four bands.	General users	6.6*3=19.8 millions images	26.5TB*3 = 79.5TB
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	218,892 tiles (same as final image atlas)	14 TB

TABLES			Rows	Volume
Preliminary Source Catalog (from coadded images)	Information on source positions, fluxes, uncertainties and SNR derived from the preliminary images (50% sky coverage)	General users	150 millions	~1 k per row
Final Science Catalog	Source catalog derived from the final science images	General users	300 millions	~1 k per row
Atlas Image Metadata	Information about each Atlas image (likely a subset of L3 Image Metadata)	General users	218,892 (one row per tile)	~1 K per row
L0 Image Metadata	Information about each Level 0 image frame	Science team only	6.6 millions (one row per image)	~1 K per row
L1 Image Metadata	Information about each Level 1b image frame	Science team only	$6.6 \times 3 = 19.8$ millions	~1 K per row
L3 Image Metadata	Information about each Level 3 image frame	Science team only	218,892 (one row per tile)	~1 K per row
L1b Source Working Database	Extracted source information from Level 1b images. 4-band, photometrically and astrometrically calibrated entries.	Science team only	~20 billions	~1 K per row
L3 Source Working Database	Extracted source information from combined Level 3 images. 4-band, photometrically and astrometrically calibrated entries.	Science team only	1 billion	~1 K per row
Calibration database	Contain all calibrators, primary & secondary.	Science team only	500	~1 K per row
Calibration table	Contain match information for all calibrators during the mission.	Science team only	10 millions	~1 K per row
Scan metadata table	Contain the final zero points for each scan	Science team only	10,000	~1 k per row
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 million	~1 K per row
Predicted Events File	List of significant events predicted to occur, including scan, time, etc.	Science team only	100,000	~ 1 K per row
Reconstructed Events File	List of significant events that actually occurred, along with actual scans, times, etc.	Science team only	100,000	~1 K per row

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 over the whole sky is expected to contain roughly 18,241 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 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 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 $\text{SNR} > 20$ in at least one band.

- ◇ The Catalog shall contain sources detected in the unconfused regions of the combined images that have $\text{SNR} > 20$ 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).
 - 1. **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 Metadata 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 purpose of developing a WISE archive loading module is to gather all of the source lists, metadata tables on the operation server disks, extracted the needed information, re-format the data when needed, and stage the data for being loaded to the IRSA database. The detailed functionality of this module is discussed as follows. IRSA will serve level 1b, level 3 and Atlas images to the public users and the WISE science team. Initially, these images will not be immediately move to the IRSA disks, and will stay on the WISE operation server disks. For IRSA to serve these fits images, we will use a table which contains the location of each fits image on the operation server disks. These addresses will be stored in image metadata tables. These tables will be generated by a WISE archive loading module, which will be developed by the members of WISE archive team. This module will extract the information related to the images from the Fits header and various other tables. Here is the list of information stored in the image metadata table for Level 1b, Level 3 and Atlas images:

◇ **Frame Image metadata table:**

- Primary positional & band information:
 1. Unique identifier (Image tile number, Scan ID, frame number)
 2. Equatorial Coordinates in J2000 (B1950, Degree and hms)
 3. Galactic coordinates
 4. Ecliptic Coordinates (computed)
 5. RA DEC of 4 corners of each tile
 6. WCS key words (put these into metadata tables)
 7. Band information

- Photometric information
 1. Photometric zero point values (from WPHOT pcal metadata tables)
 2. Noise level
 3. Background level in 4 bands --- (mean and rms, not physical units, maybe point source noise rms)

- Source detection information (frame metadata --- confusion, reliability & completeness)
 - number of sources at each band with SNR > 10?

- Time information
For Level 1b image metadata tables: we need to have
UT for each frame, associated scan ID. The time information can be taken from the image fits header: JD (Julian date), OBS_DATE, keywords:

UTCS_OBS. The associated scan information needs to get from the directory info.

- Quality, Artifact and Latency information (for both level 1b and level 3).
 1. Image QA score
 - Artifact and latency pixels
- Physical location on the operation server disks (processing metadata table)

◇ **Atlas Image (level 3 image) metadata table:**

- Primary positional & band information:
 1. Unique identifier (Image tile number, Scan ID, frame number)
 2. Equatorial Coordinates in J2000 (B1950)
 3. Galactic coordinates & Ecliptic Coordinates
 4. RA DEC of 4 corners of each tile
 5. WCS key words (put these into metadata tables)
 6. Band information
- Photometric information
 1. Photometric zero point values (from WPHOT pcal metadata tables)
 2. Noise level
 3. Background level in 4 bands --- (mean and rms, not physical units, maybe point source noise rms)
- Source detection information (frame metadata --- confusion, reliability & completeness)
 1. number of sources at each band with SNR > 10?
- Time information
--- For Level 1b image metadata tables: we need to have UT for each frame, associated scan ID. The time information can be taken from the image fits header: keywords: UTCS_OBS. The associated scan information needs to get from the directory info.
- Quality, Artifact and Latency information (for both level 1b and level 3).
 1. Image QA score
 2. Artifact and latency pixels

- Physical location on the operation server disks

◇ **Level 1b Source working database and metadata tables:**

The entire photometry tables and metadata tables from WPHOT (see System Interface Specification document

(<http://web.ipac.caltech.edu/staff/roc/wise/docs/sis/WPHot.sis>).

will be copied to IRSA (*mdex-1b.tbl, *meta-mdex-1b.tbl).

The source list will contain all detections, real sources, low SNR sources, detections of segments of large, extended objects, spurious detections, artifacts, pixel transients, latency. Each source will also have time information from frame image metadata table. These sources will have the following columns:

- **Basic Attributes:**

1. Positions, errors & epoch (J2000 equatorial in decimal degrees; also include galactic and ecliptic coordinates)
2. Fluxes & errors (4band) (should be in both flux density units and magnitude units) --- photometry should be profile fitted photometry corrected to the aperture curve of growth.
3. Flux upper limits for non-detected bands --- 2sigma or 3sigma? At the source location or the rms of the whole image?
4. Aperture photometry --- curve of growth normalized, raw aperture measurements --- how many?
5. Saturated star profile fitted.
6. Quality flags (WPHOT) ---- a) flux quality: derived from flux error, goodness-of-fit etc; b). source deblend info during profile-fitting; c). saturation: is source saturated or non-linear in one or more bands?; d). contamination: is source flux/position affected by nearby sources and/or artifacts, transients?
7. Reliability flags --- (where?) derived from SNR, detection statistics, band-detection combinations?

- **Other Desirable Attributes:**

1. Flag for extension --- use 2MASS information
2. Linkage to individual epoch of observations (how is this done?)
Flux variability (variable sources) --- to identify variable sources, compute the flux dispersion using all individual frames. This should happen either with FPG or within the level 3 source working database.

3. Position variability (solar systems & high-proper motion stars) --- with the working database, compute the position dispersion using all of the frames (multiple scans near the ecliptic pole) --- identify solar systems and high-proper motion sources.
4. Object “shape” ---- this is from profile fitting, χ^2 , derive a flag for extended source?
5. Improved photometry and positions for non-point sources - -- use aperture magnitudes?
6. Association with 2MASS --- utilize 2MASS shape and extended source flags to identify extended sources in WISE.

◇ **Level 3 Source working database and metadata tables:**

Similar to the Level 1b Source working database.

Table 3. Strawman WISE Source Catalog Record

Column	Description
ID	Unique identifier
ra, dec	J2000 (decimal deg)
errmaj,errmin, errpa	Position error ellipse (arcsec,deg EofN)
M_W1_psf,Msig_W1_psf	W1 profile-fit mag (or upper limit) and uncertainty
Repeat for W2, W3, W4	
M_W1_chi2	W1 profile-fit chi-squared
Repeat for W2, W3, W4	
M_W1_ap,Msig_W1_ap	W1 aperture mag (or upper limit) and uncertainty
Repeat for W2, W3, W4, different apertures?	
W1_cal	W1 photometric calib. (mag)
Repeat for W2, W3, W4	

Column	Description
Level 1b Working Database	
Frame_id	Identification of frameset from which this source is extracted
Scan_id	Identification of scan from which this source was extracted (this may be able to be constructed from the frame_id).
Ph_eual	Photometric quality, four character flags, one character per band. "A-best, B-next, U-upper limit"
Jul_date	Julian date correspondingn to the mid-point of the observation of the frame from which this source was detected.
sso_key	Known solar system object cross reference. Index value of entry in the known solar system object table, Null if no association
fqual	Photometric measurement quality (4 digits,1/band)
rel	Reliability flag (1 char), A-highest, B-next
cc	Contamination and confusion flag (4 char)
blend	Deblend count (int)
N/M	Num. detections and Num. of observations
Ext_flg	Extended source/shape flag (1 char). Value of character related to probability that source morphology is not consistent with single PSF
grp_id	X-ref to individual observations
Pchi2	Position distribution c^2
Fchi2_W1	W1 flux distribution c^2 (which mag?)
Repeat for W2, W3, W4	
2mass_key	Cross-reference to the "cntr/pts_key" value for closest associated 2MASS All-sky Release Point Source Catalog entry that falls within a TBD search radius of the WISE source position. Null if no associated 2MASS source
R_2mass	Radial offset between WISE source position and position of closest associated 2MASS PSC source with TBD radius (arcsec)
Pa_2mass	Position angle of separation vector from the WISE source position to the position of closest associated 2MASS PSC source within TBD radius (deg of E of N)
N2mass	The number of 2MASS All-Sky PSC entries found within a TBD search radius of the WISE source position.

[jhk]_m_2mass	J, H and Ks magnitude entries from the associated 2MASS All-Sky PSC source.
[jhk]_msig_2mass	J, H and Ks photometric uncertainties for the associated 2MASS All-Sky PSC source.
Level 3 Working Database	
Ph_qual	Photometric Quality. Four character flag, one character per band. Each character encodes the quality of the photometric measurement in the appropriate WISE band. "A" represents highest quality, "B" the next highest, etc. Special characters would include "U" for upper limit (non-detection in that band).
rel	Reliability. Single character flag. The character is related to the probability that the entry in the WDB is a real astronomical source at the time of the WISE observation. "A" is the highest probability, "B" is next highest, etc.
Var_flg	Variability flag. Single character. Character value related to the probability that the flux of this source varied in any band in an amount greater than expected from the flux measurement uncertainties.
Coadd_id	Identification of coadded image from which this source was extracted
2mass_key	Cross-reference to the "cntr/pts_key" value for closest associated 2MASS All-sky Release Point Source Catalog entry that falls within a TBD search radius of the WISE source position. Null if no associated 2MASS source
Ext_flg	Extended source flag. Single character. Character value related to the probability that the morphology of the source is not consistent with a single point spread function.
R_2mass	Radial offset between WISE source position and position of closest associated 2MASS PSC source with TBD radius (arcsec)
Pa_2mass	Position angle of separation vector from the WISE source position to the position of closest associated 2MASS PSC source within TBD radius (deg of E of N)
N2mass	The number of 2MASS All-Sky PSC entries found within a TBD search radius of the WISE source position.
[jhk]_m_2mass	J, H and Ks magnitude entries from the associated 2MASS All-Sky PSC source.
[jhk]_msig_2mass	J, H and Ks photometric uncertainties for the associated 2MASS All-Sky PSC source.

◇ **Solar systems association table:**
Talk to John (ask for 2mass example)

◇ **Scan metadata**

The archive loading module will also generate a metadata table containing information on each scan. Specifically, scan zero point values (item below) will be copied into scan metadata table. In addition, for each scan, we need to record frame number, we need to record the scan taken UT time, noise level for this scan, QA score. Do we need to store the scan processing information? (*sspipe.tbl?)

◇ **Calibration working database:**

Standards Database: column: calname, PS, NS, RA, Dec, bands, magnitudes & errors (4), spectral type, calibrator source.

Calibration Table: Columns: scan #, frame #, calname, UTC, time_from_anneal, srname, PSF_dmag & errors (4), ap_dmag & errors, quality_flag.

Scan ZP values: they are in *-spscal.tbl file, will include scan#, zp & zsig (4), time_from_anneal, #NEP, #SEP, \$PNEP, #SNEP, #PSEP, #SSEP, HC

5.2 Verification Plan and Maintenance

The WISE science center QA team will be responsible for validation of the WISE data before and after it is loaded into the archive database. Before the archive loading, the QA team will provide quality scores for various parameters, and some of these scores will be part of the data loading into the WISE archive. As discussed in Section 5.3, after the raw data are processed by the pipeline, the QA team will validate the quality of the data product. Once this verification is completed, a notice will be sent to the WISE pipeline operator to start loading the processed data into the archive database. After the loading is done, the QA team will once again verify that both source catalog and image database is running properly and serving the correct data product.

The maintenance of the archive will require collaborative work between IPAC computer support group (ISG), IRSA and WISE team. Any hardware failures will require the ISG and IRSA team to provide solutions. The WISE team and IRSA will work together to verify the integrity of the database after any such hardware failures.

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
 - Updates to match with Tim's deliveries
 - Update: August 5, 2009 (V3)
 - Update: December 30, 2009
 - Update: September 20, 2010
- **Source Query Interface Development (WISE GATOR):**
 - **Preliminary: April 8, 2009**
 - This proto-type is to test the basic structure of the archive, and the flow using the pipeline outputs from the simulated 30 orbits of data. (for example, take source tables from the wise pipeline output, run through DBprep, load into IRSA database, then allow users to query GATOR).
 - test users include IRSA and wise staff.
 - Archive QA will start planning on how to make checksums.
- **Frame Image Viewer/Access Interface:**
 - **Specification definition:** Nov 3, 2008-June 30, 2009
- **Atlas Image Viewer/Access Interface:**
 - **Specification definition:** Nov.3, 2008-June 30, 2009

5.3.1 The Integrated Schedule of IRSA-WISE Archive

Notes: Responsible party for the item is in parenthesis.

Black: IRSA delivery to WISE

Blue: WISE delivery to IRSA

Catalog Archive (GATOR) Schedule

April 15, 2009, For catalog and image meta data preparation (IRSA, done)
Spatial index code delivered to Tracey.

April 22, 2009. Catalog archive prototype (IRSA, done)

This proto-type is to test the basic structure of the archive, and the flow using the pipeline outputs from a small set of simulated data, specifically, take source tables from the wise pipeline output, run through DBprep, load into IRSA database, then allow users to query GATOR. This version will have the basic version of database dictionaries.

April 27, 2009 2MASS association code (IRSA)

Revisit the schedule

May 6, 2009 For catalog archive (IRSA)

A detailed bench mark with current hardware for loading the source list and indexing them.

A good estimate with the real hardware for loading the source list and indexing them.

Current estimates of delivery: 30 orbits in two days, 80 million rows.

May 6, 2009, For catalog archive(WISE)

Finalize the table naming convention. (ask Tracey)

May 13 - 20, 2009, For catalog archive(WISE)

- Large simulated data(in observation pattern) to test the fast loading and index speed

- Table names follow the agreed upon naming convention

- About 2 billion rows of source list with spatial index generated

May 20, 2009, For catalog archive (IRSA)

- 2MASS association code

- Bench mark the time needed to generate the XPF files

June 10, 2009 For catalog archive V1 (WISE)

- Tracey's DBprep v2.8

- Large simulated data sets from WISE for V1 release.

- Include the spacial index and 2MASS association.
- Data dictionary updated
- Those data table files can be delivered one at a time. We need time to work out the issues.
- Release audience: WISE team

-

June 17, 2009 For catalog archive(IRSA)

- All the data tables loaded into the IRSA database for WISE
- XPF files generated

June 24, 2009 For catalog archive (IRSA)

- GATOR interface ready for IPAC WISE team to check out

July 1, 2009 Catalog archive V1 (IRSA)

- GATOR interface ready for WISE team

August 3, 2009 For catalog archive (WISE)

- DBprep + dictionary v3 update
- Updated database tables and data dictionaries

Sept. 9, 2009 For catalog archive V2(WISE)

1. Final pre-launch update on database dictionary
2. Table updates

Sept. 30, 2009 For catalog archive (WISE)

1. When is the first source list available from WISE operation? Nov. 1, 2009
2. What is the frequency and data volume to deliver to IRSA for ingest? Twice/wk, 80 million rows per load (3 days)
3. What is the expected turn around from delivery to accessible to WISE team? 3days

Oct. 6, 2009 Catalog archive V2 (IRSA)

1. Updated data dictionary
2. Updated tables

Jan. 16, 2010 For catalog archive (WISE)

- DBprep + dictionary v3.5 update
- Updated database tables and data dictionaries

Feb 10, 2010 Catalog archiveV3 (IRSA)

- After-launch & after IOC updates

Feb. 15, 2010 For catalog archive (tentative WISE)

- [Requirements on solar system object search](#)

Sept. 10, 2010, Catalog archive V4 (tentative IRSA)

- 1) Solar system object interface (object name; user supplied orbital info --- Image archive)

Image Archive Schedule

April 22, 2009, For image archive (WISE)

[Schema for L1B and L3 image meta data](#)

1. All the information listed in Lin's Archive Design documents
2. x, y, z, spt_idx (Contact Tracey)

April 27, 2009 2MASS association code (IRSA)

Revisit the schedule

May 6, 2009. Image archive plan (IRSA)

A detailed implementation plan and schedule for the Web interface.

May 6, 2009 Look and Feel for image access (IRSA)

HTML mock-up of the wise image archive web interface (input and output pages)

May 15, 2009, For image archive (WISE)

- Test data: Meta data and image files consistent with meta data for L1b and L3 . (Tracey)
- Algorithm for 3 color image generation (Davy)
- One default stretch (Davy)
- Define what is the "best" frame set (Lin)
Position most centered.

June 3, 2009, Image archive prototype (IRSA)

The basic software structure in place.

For L1b and L3/atlas images:

1. Position/fixed object name search
2. ScanID + FrameNumber
3. WISESourceID

Default output: "best" frame sets (WISE will define "best")

Optional Input:

- Date (date only for proto-type)
- Band selection (default input to all)
- Image size (smaller than original)

Output:

List of image metadata

Single frame set image display, native pixel, 3 color JPEG where applicable (Only a simple algorithm such as weighted averaging)

Allow download FITS images, including noise and mask images for L1b, noise and depth of coverage for L3

No image mosaic, no re-projection

June 30, 2009 For image archive (WISE)

- Feed backs for image archive prototype
- Limit/throttle control without impact the operational system since IRSA will access the files through NFS mount. (Tim)

July 15, 2009, Image archive prototype 2 for internal demo (IRSA)

- Overlay artifacts (default). Allow optional overlay the source list
- User can adjust stretches of displayed image

Aug. 10, 2009 For image archive (WISE)

- The image meta data tables for L1b and L3 data
- The data products files associated with the meta data

Sept. 9, 2009 Image archive V1 (IRSA)

- Image inventory
- Batch search: wget scripts for the final public releases.
- Allow search with date range
- Real delivery of overlay features

Multiple epoch images display, no re-projection with download volume control (need to define this volume, WISE)

[Add dss, 2mass images display with L3/atlas images: this could be included in FindingCharts]

Nov. 18, 2009 Image archive V2 (IRSA)

- Polish and bug fixes
- Multiple epoch images display with re-projection for L3/atlas images only
- Download re-projected images

July 15, 2010 Image archive V3 (IRSA)

- L3/Atlas data
- Arbitrary User input pixel size, reprojection and image mosaic
- Re-project, re-sample pixel scale, and make mosaic on-fly.
- Always return a square image output
- Add dss, 2mass images display with L3/atlas images: this could be included in FindingCharts

Sept. 10, 2010 Image archive V4 (IRSA)

Solar system:

Name (known solar object name)

Skylocation/Time (or new source)

6 STAFFING PLAN

Currently, the people from IRSA working on the WISE data archive include Angela Zhang, Serge Monkewitz, Ramon Rey and Steve Groom. Angela and Ramon are primarily responsible for the GATOR

Interface and source database loading. Serge will develop the image archive.

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.

