LARGE SYNOPTIC SURVEY TELESCOPE

Large Synoptic Survey Telescope (LSST)

Data Management Releases for Verification/Integration

William O'Mullane, Frossie Economou, Tim Jenness, Andrew Loftus, John D. Swinbank

LDM-564

Latest Revision: 2018-04-13

Abstract

This document describes release management at a high level and specific features for upcoming releases.

Change Record

Version	Date	Description	Owner name
1.0	2017-08-18	Initial version. Approved in RFC-373.	W. O'Mullane
1.1	2018-03-16	Synchronize milestones with PMCS.	J.D. Swinbank

Document source location: `https://github.com/lsst/LDM-564`

Contents

Data Management Releases for Verification/Integration

1 Introduction

1.1 Scope

This document covers releases of software from the Data Management Subsystem of LSST for verification/integration tests. It discusses the delineation between the Data Facility as an operational entity and DM producing and testing software. It does not cover the normal releases to the community of the software stack - see `https://developer.lsst.io/` for that.

2 Release Management

This section outlines the current understanding of the release management process. Complete definition is pending the appointment of the DM Release Manager.

2.1 Preparation of Releases

DM develops code in GitHub following its developer guidelines and coding standards [1]. This includes automated testing and continuous integration. Tested releases are tagged by SQuaRE weekly and major releases are made each cycle (six months).

There are specific packages and systems deployed together to form the high level components of DM as depicted in Figure 1. The orchestration of deployments on multiple machines is facilitated by the use of containers and machine readable configurations. DM prepares Docker containers and Puppet configurations for deploying these systems on Kubernetes enabled clusters. These artifacts are tagged as part of the release.

In addition, specific releases with features required to support the LDM-503 test milestones will be tagged and released in advanced of each verification test. The preliminary feature lists for these releases are defined in Section 3.

[1] `https://developer.lsst.io/`

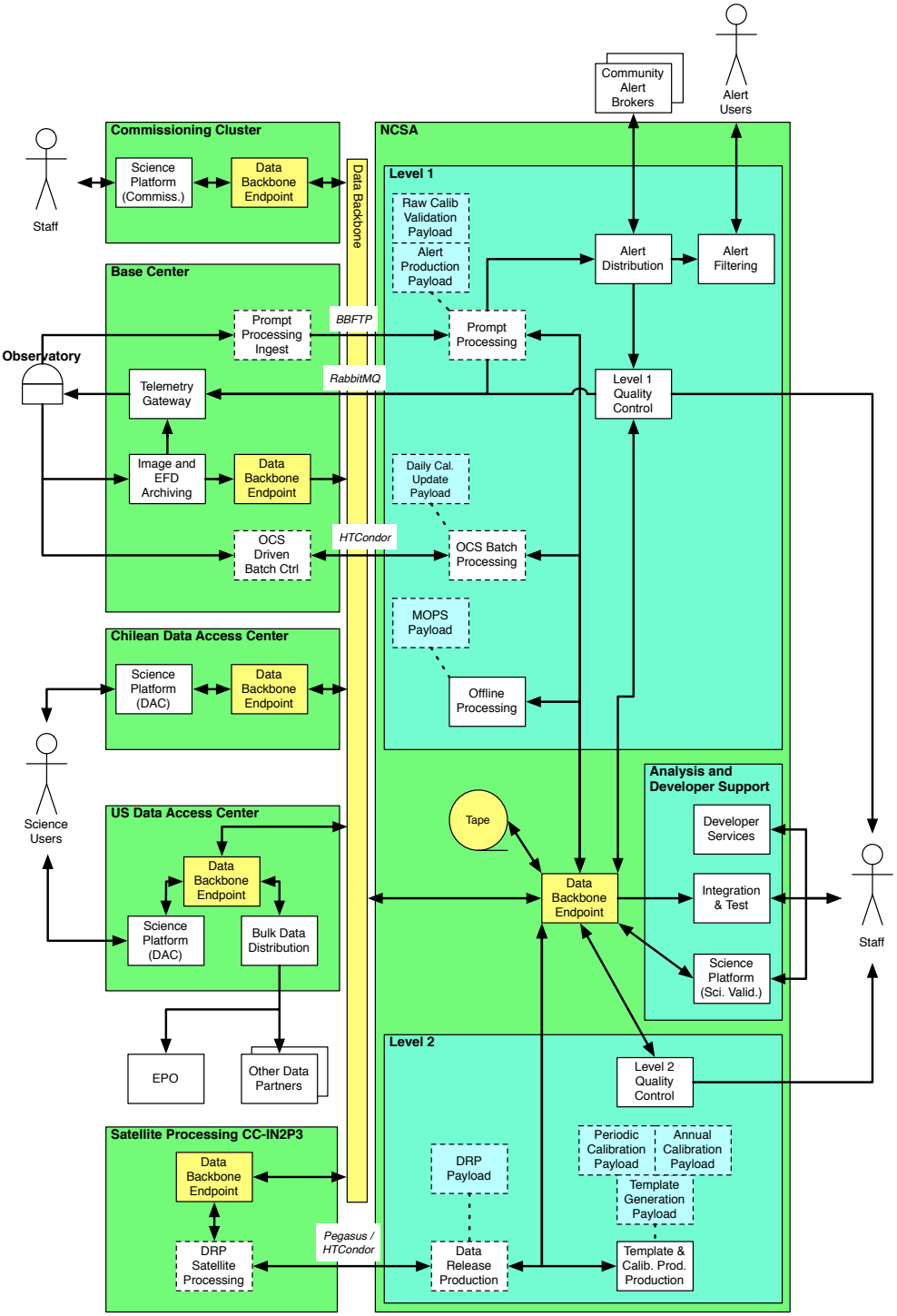

FIGURE 1: DM components as deployed during Operations. Where components are deployed in multiple locations, the connections between them are labeled with the relevant communication protocols. Science payloads are shown in blue. For details, refer to LDM-148.

2.2 Deployment of Releases

Although DM will provide ready-to-install products, these will be further tested before being deployed. Hence, releases will initially be installed on test systems at NCSA and will undergo smoke testing before they are made available in the production environment. This will serve as an operational validation of the release.

Once smoke tested, the Docker containers will be made available in the NCSA Docker repository. Using this secure internal repository, operators may deploy containers for specific releases in the operational environment.

2.2.1 Levels of operational validation

Certain containers will be used to provide kernels and supporting libraries for the JupyterLab environment. Multiple versions of these containers can be made available simultaneously — for example, providing a series of minor releases of the software stack — with the user selecting which to deploy for their particular use case. Since they will not be deployed as part of the core operational system, acceptance testing can be relatively minimal.

Some containers will be made available on development systems in support of ongoing development of the code. Again, these should be made available rapidly, with security checking and validation testing kept to a minimum.

Similarly, during Commissioning, availability of containers on the Commissioning Cluster should be on the order of hours (not days). The level of smoke testing and the time to availability of a container may need to be compressed in Commissioning.

Containers to be used for prompt or batch processing on operational systems, on the other hand, must be rigorously validated.

3 Functionality in future DM releases

This is currently not an exhaustive feature list, but rather gives an indication at a high level of the features in each release which will be verified by the corresponding verification test campaign. As the test plans are written this will become a list of requirements to be tested for that release and thus begin to fill out the verification control database (currently to be in JIRA).

In the feature lists below, the corresponding internal milestone is given in parenthesis.

Each section here is a test milestone from LDM-503 — the same labels are used. The timeline is in the DM schedule using the same labels and depicted in Figure 2

3.1 Science Platform with WISE data in PDAC: LDM-503-01

Due: 2017-11-30; currently incomplete.

- DM-DAX-2: Query service supporting IVOA TAP protocol, w/ support for asynchronous queries *(Due: 2017-07-31; currently incomplete)*
- DM-SQRE-1: Project internal Jupyter notebook service *(Due: 2017-08-31; completed 2017-11-01)*
- DM-DAX-4: Metadata service supporting IVOA SIAv2 protocol *(Due: 2017-09-29; currently incomplete)*
- DM-DAX-1: WISE data ingest to PDAC *(Due: 2017-11-30; completed 2017-11-01)*
- DM-DAX-3: Image cutout service supporting IVOA SODA protocol *(Due: 2017-11-30; currently incomplete)*

3.2 HSC reprocessing: LDM-503-02

Due: 2017-11-30; completed 2017-12-01.

- DM-DRP-1: HSC merger complete: all functionality deployed for the most recent HSC data release processing is now available within the LSST stack. *(Due: 2017-05-31; completed 2017-11-01)*
- DM-NCSA-1: Provide regular reprocessing service for HSC data *(Due: 2017-05-31; completed 2017-11-01)*
- DM-NCSA-2: Provide access to results of regular reprocessing (NB the form this takes depends upon available DAX functionality) *(Due: 2017-05-31; completed 2017-11-01)*
- DM-AP-1: Basic single frame measurement pipeline. *(Due: 2017-08-31; completed 2017-11-01)*

FIGURE 2: DM level 2 milestones (LDM-503-x) in the LSST schedule.

- DM-DAX-5: Database ingest in support of HSC reprocessing (ie, large catalog ingest) *(Due: 2017-11-30; currently incomplete)*
- DM-DRP-2: Basic visualization and quality assessment tools operational on HSC-scale data volumes. *(Due: 2017-11-30; completed 2017-11-01)*
- DM-NCSA-3: Provide database for metadata, provenance, location and demonstrate ingest at small scale *(Due: 2017-11-30; completed 2017-11-01)*

3.3 Alert generation validation: LDM-503-03

Due: 2017-11-30; completed 2017-12-01.

- DM-AP-1: Basic single frame measurement pipeline. *(Due: 2017-08-31; completed 2017-11-01)*
- DM-AP-2: Alard & Lupton-style image differencing. *(Due: 2017-08-31; completed 2017-11-01)*
- DM-AP-3: Point source & dipole measurement on difference images. *(Due: 2017-08-31; completed 2017-11-01)*
- DM-AP-4: DIASource association *(Due: 2017-08-31; completed 2017-11-01)*
- DM-AP-5: DIAObject generation *(Due: 2017-08-31; completed 2017-11-01)*
- DM-DAX-6: Prototype level 1 database *(Due: 2017-11-30; completed 2017-11-01)*

3.4 Aux Tel DAQ integration functionality test: LDM-503-04

Due: 2018-06-01; currently incomplete.

- DM-NCSA-4: Minimal support for the small operational schema including file metadata and provenance for every file, and record of in *(Due: 2017-11-30; currently incomplete)*
- DM-DAX-7: Butler interface to retrieve images from data backbone *(Due: 2018-06-01; currently incomplete)*

3.5 Aux Tel DAQ interface Integration Verification and Spectrograph Operations Rehearsal: LDM-503-04b

Due: 2018-06-01; currently incomplete.

- DM-NCSA-6: Ability to transfer files originating from Tucson to NCSA and ingest files at NCSA, including metadata and provenance *(Due: 2018-03-05; currently incomplete)*
- DM-NCSA-5: Level 1 archiving system able to acquire pixel data from the Aux Tel DAQ, header metadata via OCS, assemble FITS image, *(Due: 2018-03-30; currently incomplete)*
- DM-NCSA-7: Capability to paint displays for Tucson and NCSA *(Due: 2018-03-30; currently incomplete)*

3.6 Alert distribution validation: LDM-503-05

Due: 2018-07-02; currently incomplete.

- DM-NCSA-8: Test instance of feeds to LSST mini broker in online (live stream) and offline (replaying from files) modes *(Due: 2018-06-01; currently incomplete)*
- DM-NCSA-9: Test instance of alert distribution hosting service and L1 database in Development & Integration Enclave *(Due: 2018-06-01; currently incomplete)*
- DM-AP-6: Alert format defined & queue system available. *(Due: 2018-07-02; currently incomplete)*

3.7 Small Scale CCOB Data Access: LDM-503-8b

Due: 2018-09-06; currently incomplete.

3.8 DM ComCam interface verification readiness: LDM-503-06

Due: 2018-11-30; currently incomplete.

- DM-DRP-4: Calibration product generation in support of basic ISR. *(Due: 2017-05-31; completed 2017-12-01)*

- DM-NCSA-10: Sustained archiving service that is OCS commandable *(Due: 2018-06-22; currently incomplete)*
- DM-AP-7: Basic instrument signature removal (ISR) capability. *(Due: 2018-06-29; currently incomplete)*
- DM-DRP-3: PSF-homogenized coadd construction. *(Due: 2018-06-29; completed 2017-11-01)*
- DM-DRP-38: Camera package supporting the Commissioning Camera. *(Due: 2018-06-29; currently incomplete)*
- DM-NCSA-11: Verified acquisition of raw and crosstalk-corrected exposures at raft scale, incl. correct metadata *(Due: 2018-11-30; currently incomplete)*

3.9 Camera data processing: LDM-503-07

Due: 2018-11-30; currently incomplete.

- DM-DRP-5: Camera package supporting the LSST Camera. *(Due: 2018-06-29; currently incomplete)*

3.10 Spectrograph data acquisition: LDM-503-08

Due: 2018-11-30; currently incomplete.

- DM-DRP-6: Camera package supporting the Auxiliary Telescope. *(Due: 2017-08-31; currently incomplete)*
- DM-NET-6: Summit LAN installed *(Due: 2018-03-05; currently incomplete)*
- DM-DRP-7: Coordinate transformation tool provided for use with the Collimated Beam Projector. *(Due: 2018-06-01; currently incomplete)*
- DM-AP-7: Basic instrument signature removal (ISR) capability. *(Due: 2018-06-29; currently incomplete)*
- DM-NCSA-12: EFD ETL Service *(Due: 2018-09-28; currently incomplete)*
- DM-NCSA-14: Data Backbone endpoints in Chile for ingestion and access, distribution over WAN, ingest at NCSA into custodial file sto *(Due: 2018-09-28; currently incomplete)*

- DM-DRP-8: Calibration product generation for the Auxiliary Telescope. *(Due: 2018-11-30; currently incomplete)*
- DM-DRP-9: Data reduction pipeline for the Auxiliary Telescope. *(Due: 2018-11-30; currently incomplete)*
- DM-NCSA-13: Header Writing Service for Spectrograph use case *(Due: 2018-11-30; currently incomplete)*
- DM-NCSA-15: Batch Processing Service for offline spectrograph data processing *(Due: 2018-11-30; currently incomplete)*

3.11 DAQ validation: LDM-503-10

Due: 2018-11-30; currently incomplete.

3.12 Ops rehearsal for commissioning #1: LDM-503-09

Due: 2019-01-04; currently incomplete.

- DM-DRP-14: Insertion of simulated sources into the data stream to check pipeline performance. *(Due: 2017-11-30; completed 2017-12-01)*
- DM-DRP-18: Initial multi-band deblending algorithm available. *(Due: 2017-11-30; completed 2017-12-01)*
- DM-DRP-16: Global photometric fitting (e.g. Burke et al. Forward Global Calibration Method). *(Due: 2018-01-31; currently incomplete)*
- DM-DRP-11: Pipelines code provides supports for database ingestion of results. *(Due: 2018-05-31; currently incomplete)*
- DM-AP-9: Jointcal at a functional (but not necessarily algorithmically complete) level. *(Due: 2018-07-20; currently incomplete)*
- DM-DRP-17: Simultaneous photometric and astrometric fitting to multiple exposures. *(Due: 2018-07-20; currently incomplete)*
- DM-DAX-8: Supertask-based system capable of efficient processing across a full focal plane. *(Due: 2018-11-30; currently incomplete)*

- DM-DAX-9: Provenance system, details TBD. *(Due: 2018-11-30; currently incomplete)*
- DM-DRP-10: Calibration products include an optical ghost model. *(Due: 2018-11-30; currently incomplete)*
- DM-DRP-12: Background estimation over the full visit. *(Due: 2018-11-30; currently incomplete)*
- DM-DRP-13: PSF estimation over the full visit. *(Due: 2018-11-30; currently incomplete)*
- DM-DRP-15: All varieties of coadd required for object detection and characterization are now produced during normal pipeline operation (although not necessarily at the ultimately required level of fidelity). *(Due: 2018-11-30; currently incomplete)*
- DM-DRP-19: QA metrics are generated during pipeline execution. *(Due: 2018-11-30; currently incomplete)*
- DM-DRP-37: Artifact rejection and background matching during coadd construction. *(Due: 2018-11-30; currently incomplete)*
- DM-NCSA-16: Perform ISR processing on ComCam-scale data. *(Due: 2018-11-30; currently incomplete)*
- DM-NCSA-17: QA on WCS, PSF, etc returned to Observatory using JupyterLab *(Due: 2018-11-30; currently incomplete)*
- DM-NCSA-18: Validated disaster repsonse recovery for data and calibration products *(Due: 2018-11-30; currently incomplete)*
- DM-NCSA-19: 8x7 incident response system *(Due: 2018-11-30; currently incomplete)*
- DM-SQRE-2: Commissioning notebooks running on the commissioning cluster *(Due: 2018-11-30; currently incomplete)*
- DM-AP-8: Advanced ISR, including ghost and linear feature masking, correction for the Brighter-Fatter effect and compensation for pixel response non-uniformity. *(Due: 2019-01-04; currently incomplete)*

3.13 Large Scale CCOB Data Access: LDM-503-10b

Due: 2019-08-06; currently incomplete.

3.14 DM ComCam operations readiness: LDM-503-11a

Due: 2019-09-30; currently incomplete.

- DM-AP-2: Alard & Lupton-style image differencing. *(Due: 2017-08-31; completed 2017-11-01)*
- DM-AP-3: Point source & dipole measurement on difference images. *(Due: 2017-08-31; completed 2017-11-01)*
- DM-AP-6: Alert format defined & queue system available. *(Due: 2018-07-02; currently incomplete)*
- DM-AP-10: Advanced single frame measurement pipeline for Alert Production. *(Due: 2019-08-20; currently incomplete)*
- DM-NCSA-20: ComCam Archiving Service *(Due: 2019-09-30; currently incomplete)*

3.15 Ops rehearsal for commissioning #2: LDM-503-11

Due: 2019-10-31; currently incomplete.

- DM-DAX-10: Middleware support for multifit *(Due: 2018-11-30; currently incomplete)*
- DM-AP-11: Difference imaging includes noise decorrelation and correction for differential chromatic refraction. *(Due: 2019-01-08; currently incomplete)*
- DM-DRP-21: Integrated image characterization pipeline for Data Release Production. *(Due: 2019-05-31; currently incomplete)*
- DM-DRP-22: Template generation integrated with Data Release Production pipelines. *(Due: 2019-05-31; currently incomplete)*
- DM-DRP-23: Atmospheric characterization based on data from the Auxiliary Telescope now available. *(Due: 2019-05-31; currently incomplete)*
- DM-DRP-26: Overlap resolution at tract & patch boundaries. *(Due: 2019-10-31; currently incomplete)*

3.16 Ops rehearsal for commissioning #3: LDM-503-12

Due: 2020-01-31; currently incomplete.

- DM-DRP-29: Moving point source model fitting now available. *(Due: 2018-11-30; currently incomplete)*
- DM-DRP-20: Refined set of LSST calibration products. *(Due: 2019-05-31; currently incomplete)*
- DM-SQRE-3: Hardened Jupyter deployment on Commissioning Cluster *(Due: 2019-08-05; currently incomplete)*
- DM-DRP-24: Physically motivated PSF model, including separate characterization of contributions from the atmosphere and the telescope system. *(Due: 2019-11-27; currently incomplete)*
- DM-DRP-27: Object generation: association and assembly of (DIA, coadd, etc) sources to form objects. *(Due: 2019-11-27; currently incomplete)*
- DM-DRP-28: Difference images are now a first-class data product during data release processing. *(Due: 2020-01-31; currently incomplete)*
- DM-DRP-30: Forced photometry is now performed on individual processed visit images during data releases. *(Due: 2020-01-31; currently incomplete)*

3.17 Ops rehearsal DRP (ComCam data): LDM-503-13

Due: 2020-11-30; currently incomplete.

- DM-DRP-32: Object classification system available. *(Due: 2018-03-30; currently incomplete)*
- DM-DRP-33: Generation of coadded images suitable for use in EPO activities. *(Due: 2018-11-30; currently incomplete)*
- DM-AP-12: Difference imaging is now agnostic to the PSF of the template image. *(Due: 2019-04-24; currently incomplete)*
- DM-STAFF: Staffing Checkpoint *(Due: 2019-12-02; currently incomplete)*

- DM-AP-13: Trailed source measurement on difference images. *(Due: 2020-03-25; currently incomplete)*
- DM-DRP-25: Prototype multi-epoch fitting system available. *(Due: 2020-05-29; currently incomplete)*
- DM-DRP-34: Selection maps are generated during data releases. *(Due: 2020-05-29; currently incomplete)*
- DM-AP-14: Alert filtering system available. *(Due: 2020-06-15; currently incomplete)*
- DM-DRP-35: Simultaneous measurement across a suite of coadds representing different bandpasses, epocs, and flavors. *(Due: 2020-11-30; currently incomplete)*
- DM-NCSA-23: Operational processes for preparing for & producing a data release defined and tested *(Due: 2020-11-30; currently incomplete)*

3.18 DM Software for Science Verification: LDM-503-14

Due: 2020-11-30; currently incomplete.

- DM-DRP-33: Generation of coadded images suitable for use in EPO activities. *(Due: 2018-11-30; currently incomplete)*
- DM-AP-15: Alert distribution system fully integrated. *(Due: 2019-10-09; currently incomplete)*
- DM-SQRE-4: Notebook service ready for verification & validation *(Due: 2020-02-10; currently incomplete)*

3.19 Full scale ops rehearsal: LDM-503-15

Due: 2021-10-29; currently incomplete.

- DM-SQRE-5: Notebook service ready for general science users *(Due: 2020-08-07; currently incomplete)*
- DM-NCSA-25: Demonstrate operational coordination with and processing at satellite CC-IN2P3 satellite computing facility *(Due: 2020-09-30; currently incomplete)*

- DM-NCSA-24: Production batch service for data release processing *(Due: 2020-11-30; currently incomplete)*
- DM-DRP-31: A photometric redshift is now provided for each object. *(Due: 2021-05-31; currently incomplete)*
- DM-AP-17: Moving object processing system (MOPS) available. *(Due: 2021-08-12; currently incomplete)*
- DM-AP-16: Full integration of the Alert Production system within the operational environment. *(Due: 2021-09-03; currently incomplete)*
- DM-DRP-36: MOPS integrated to data release processing. *(Due: 2021-10-04; currently incomplete)*

3.20 Ops rehearsal DRP: LDM-503-16

Due: 2022-05-31; currently incomplete.

- DM-NCSA-26: Demonstrate full delivery of Data Facility ConOps *(Due: 2022-05-31; currently incomplete)*

3.21 Verification Tests for Full DRP: LDM-503-17

Due: 2022-05-31; currently incomplete.

4 References

[1] **[LDM-148]**, Lim, K.T., Bosch, J., Dubois-Felsmann, G., et al., 2017, *Data Management System Design*, LDM-148, URL `https://ls.st/LDM-148`

[2] **[LDM-503]**, O'Mullane, W., Jurić, M., Economou, F., 2017, *Data Management Test Plan*, LDM-503, URL `https://ls.st/LDM-503`

5 Acronyms

Acronym	Description
DAQ	Data AcQuisition (system)
DAX	Data access services
DM	Data Management
DRP	Data Release Production
EFD	Engineering Facilities Database
FITS	Flexible Image Transport System
ISR	Instrument Signal Removal
IVOA	International Virtual-Observatory Alliance
JIRA	issue tracking product (not an acronym, but a truncation of Gojira, the Japanese name for Godzilla)
LAN	Local Area Network
LSST	Large Synoptic Survey Telescope
MOPS	Moving Object Pipelines
NCSA	National Center for Supercomputing Applications
OCS	Observatory Control System
PDAC	Prototype Data Access Center
PMCS	Project Management Control System
PSF	Point Spread Function
QA	Quality Assurance
SODA	SCOS ORATOS Distributed Access
SQuaRE	Science Quality and Reliability Engineering
TAP	Table Access Protocol
TBD	To Be Defined (Determined)
WCS	World Coordinate System
WISE	Wide-field Survey Explorer