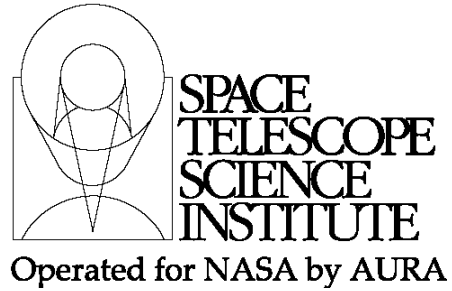




# TECHNICAL REPORT



<b>Title: JWST I&amp;T Data Archive Operations Concept</b>	Doc #: JWST-STScI-001658, SM-12 Date: February 2, 2009 Rev: -
Authors: Rusty Whitman Phone: 410-338-5042	Release Date: February 13, 2009

## 1.0 Introduction

The primary purpose of the JWST I&T Data Archive is to provide a permanent repository for selected I&T data files to support pre-launch generation of the initial set of instrument calibration reference files and to provide a reference archive of I&T data for post-launch anomaly investigation. Secondly, the archive will provide a central location for instrument teams to pool and share data, and provide a source of test data files for S&OC and ground segment testing. The I&T Data Archive is a component of the S&OC Data Management System.

## 2.0 I&T Data Ingest Concept

There will be two types of ingest into the I&T Data Archive; package based, and file based (for FITS files only).

Package based ingest will be used to ingest I&T Data Packages. An I&T Data Package is a directory structure containing data files from a specific I&T activity. The files placed in the directory structure may be of any type (text, FITS, SSR, Zip, etc). A Readme.xml file in the top level directory of the package will contain the metadata that describes the package and will be used to populate the catalog entry for the package. On ingest this entire directory structure will be gathered into a tar file, assigned a unique name, and ingested into the I&T Data Archive. The unit of ingest and retrieval for an I&T Data Package is the tar file. Every package submission is cataloged as a new entry in the archive. To associate a package with a previously submitted package the keywords in the Readme.xml are set appropriately (see example below).

File based ingest will be used to ingest I&T science FITS files that are provided to STScI by the SI teams to support the generation of the initial set of calibration reference files. A file based submission consists of a set of uniquely named science FITS files in a directory. Each FITS file is cataloged and archived separately. A Readme.xml file in the

**Operated by the Association of Universities for Research in Astronomy, Inc., for the National Aeronautics and Space Administration under Contract NAS5-03127**

Check with the JWST SOCCER Database at: <http://soccer.stsci.edu/DmsProdAgile/PLMServlet>  
To verify that this is the current version.

same directory as the FITS files will contain the metadata that describes the submission and that will be used to populate the common catalog entries for the submission. Non-FITS files, with the exception of the Readme.xml file, are ignored. Selected FITS keywords from the FITS files will be extracted and also added to the catalog. The document “JWST Ground Test Data Standard FITS Keywords”, JWST-STScI-001487 provides guidelines for additional FITS keywords that should be included in the files, however only the FITS keywords listed below will become part of the archive catalog. Similar to package based submissions, the FITS files in the file based submission can be associated by setting the keywords in the Readme.xml file appropriately. Unlike package based submissions the submission of a FITS file that has the exact filename as a previously cataloged FITS file will result in the newer file replacing the older file. This allows for the submission of “best” versions of the I&T science FITS files. On query and retrieval only the “best” version is presented to the user. See the examples section below for samples of file based submissions.

The Readme.xml file will contain the following information:

- Area – Facility where test is being conducted
- Phase – Ambient, Cryo, etc
- Start date and time – Date and time of start of test
- End date and time – Date and time of end of test
- Test title – As specified in test plan
- Test number – As specified in test plan
- Responsible organization – Organization conducting the test
- Responsible engineer – Point of contact for this test
- Test description – Short text description of purpose/content of this test
- Package Type – ‘Package’ or ‘File’

A cecil proc will be developed by **TBD** that will gather the above information and create the Readme.xml file for a test. The cecil proc will be run by the test conductor who will be responsible for gathering the data files for an I&T Data Package and providing the package to the STScI for ingest. The Readme.xml file required for file based submissions of science FITS files should in general be a modified version of the Readme.xml file from the test that generated the input science data packetized files. The cecil proc should be used whenever possible to generate the Readme.xml file as it will be coded to ensure consistency in the values of the keywords.

The FITS keywords that are required to be present in the science FITS files for ingest into the I&T Data Archive are populated by the FITSWriter in the Level 1 FITS files. Higher level processing of the Level 1 science files should retain these keywords and update the values if necessary. The keywords extracted from the FITS header and populated in the catalog appear in the table below. Note that the keyword TARGNAME will be initially populated by the FITSWriter (default is UNKNOWN) but that the user is responsible for updating the keyword with a meaningful value.

Check with the JWST SOCCER Database at: <http://soccer.stsci.edu/DmsProdAgile/PLMServlet>  
To verify that this is the current version.

Keyword	Description	Source
DATE	date this file created(yyyy-mm-ddThh:mm:ss, UTC)	System clock
ROOTNAME	rootname of the observation set	Program generated
INSTRUME	identifier for instrument used to acquire data	Telemetry
OBS_ID	Observation ID from the science data packet image header	Telemetry
DATAMODE	Reference number used to specify ground data processing requirements for this data	Telemetry
TARGNAME	Name of target	User supplied
EXPTIME	Exposure time for observation	Telemetry derived

### 3.0 Ingest Examples

#### 3.1 I&T Data Package Submission

- 1) The test conductor completes a setup procedure prior to running the test. This involves creating the directory structure to hold the results of the test and running a cecil proc that creates the Readme.xml file for the test. The test conductor may edit the Readme.xml file as needed.
- 2) At the conclusion of the test the test conductor collects the data products to be sent to the I&T Data Archive and puts them into the results directory structure. The Readme.xml file is put into the root directory. This directory structure and contents constitutes the data package for archival.
- 3) The test conductor notifies the S&OC of the location of the Data Package and requests that the Data Package be archived. Nominally the data package will be placed on a volume that is mirrored to a disk volume local to the S&OC, however Data Packages may also be made available via the network or provided on media shipped to the S&OC.
- 4) S&OC personnel run the ingest process to ingest the I&T Data Package.
  - a. The ingest process reads the Readme.xml file.
  - b. The ingest process creates a unique filename for the I&T Data Package and generates a tarball of the entire directory structure.
  - c. The tarball is archived and a catalog entry is created using the information from the Readme.xml file. In addition a field is set to indicate that this is a I&T Data Package.
- 5) S&OC personnel notify the test conductor that the I&T Data Package was successfully archived.

Note: In the event that a Data Package must be broken into multiple pieces for submission (e.g. due to size or test conductor preference) or that a test conductor wishes to submit additional files to be added to a previous submission, then the Readme.xml file parameter values for each submission must be the same with the exception of the Description field. This will ensure that the multiple Data Packages are all associated in the catalog.

### **3.2 I&T Data Package**

#### **3.3 Individual FITS File Submission**

- 1) Instrument teams will process selected science data files into higher level FITS products. These FITS files are put into a directory along with a Readme.xml file and delivered to the S&OC. This constitutes a File-Based package.
- 2) S&OC personnel run the ingest process to ingest the File-Based Package.
  - a. The ingest process reads the Readme.xml file and extracts the keywords from the file.
  - b. For each FITS file in the directory:
    - i. Each FITS file is verified by running it through fitsverify.
    - ii. The ingest process reads the required FITS header keywords
    - iii. The ingest process archives the FITS file.
    - iv. The ingest process creates a catalog entry for the FITS files using the keywords from the Readme.xml file and the FITS header keywords. If the filename for this entry is a duplicate of the filename for an existing entry then the new entry replaces the previous entry.
- 3) S&OC personnel notify the POC for the instrument team that the File-Based Package was successfully archived.

### **4.0 Query/Distribution Examples**

#### **4.1 I&T Data Package Example**

- 1) Users access MAST and select the JWST I&T Archive query screen.
- 2) To find all I&T Data Packages associated with a specific Test and Test Number the user selects the I&T Data Packages only radio button (otherwise FITS files matching the query will also be returned), enters the Test name and Test Number into the appropriate fields and clicks on Submit (selection lists are used wherever possible).
- 3) A list of Data Packages is presented to the user along with metadata associated with each of the entries.
- 4) To request a Data Package from the archive the user selects the desired entries and clicks on Submit Retrieval Request.
- 5) The user is prompted to login to verify they are authorized to retrieve the requested data.
- 6) The user specifies the delivery method for the requested data.

#### **4.2 Individual FITS File Example**

- 1) Users access MAST and select the JWST I&T Archive query screen.
- 2) To find all of the NIRCcam DARKs taken using a readout pattern of DEEP8 the user selects NIRCcam from the instrument field, DEEP8 for Datamode (a cross reference will be provided between the datamode number and name), and 'DARK' for the Target Name.
- 3) A list of FITS files is presented to the user along with metadata associated with each of the entries.

Check with the JWST SOCCER Database at: <http://soccer.stsci.edu/DmsProdAgile/PLMServlet>  
To verify that this is the current version.

- 4) To request FITS files from the archive the user selects the desired entries and clicks on Submit Retrieval Request.
- 5) The user is prompted to login to verify they are authorized to retrieve the requested data.
- 6) The user specifies the delivery method for the requested data.

Check with the JWST SOCCER Database at: <http://soccer.stsci.edu/DmsProdAgile/PLMServlet>  
To verify that this is the current version.