

SID Archive Cookbook

January 29, 2015

Table of Contents

Introduction.....	2
Nomenclature.....	2
Preparing Data for Archiving.....	3
Determine the Submission Type.....	3
Create the Readme.xml File	4
Helpful Hints for Manual Creation of the Readme.xml file	7
Verify the FITS files	7
FITS Header Keyword ROOTNAME.....	9
Delivering Data to the Archive.....	9
Dataset Naming Conventions in the SID Archive	10
SID Archive ITAR Handling	10
Searching for Data in the SID Archive	10
Retrieving Data	12
Proprietary Access	17
Appendix A Readme.xml schema	18

Introduction

The SID archive is a repository of pre-launch JWST instrument test data. It is intended to be a convenient place for teams to store data and from which data may be retrieved. In particular, I&T data from the *individual* science instruments are needed to support a variety of activities. These include

- Support generation of initial instrument calibration reference files
- Provide source of test data for ground segment I&T
- Provide an historical data archive of instrument I&T data for use in anomaly resolution for SI Operations
- Support collaborative efforts between SI teams

In addition, I&T data from *integrated* Observatory tests need to be captured and stored at STScI to support the following activities.

- Provide source of test data for ground segment I&T
- Provide an historical data archive of I&T data for use in anomaly resolution for Observatory Operations

In this cookbook, there is an outline of the procedure to follow when sending test data to the archive. There is also a discussion of the archive interface, which is used to search the archive for data, and a brief description of the Data Archive and Distribution System (DADS), which distributes data from the archive. Examples are included.

All data in the SID archive are proprietary. Since these data will never become public, anonymous data retrieval will not be possible.

Nomenclature

There are a handful of important terms that appear in this cookbook. The terms are listed and discussed here.

Submission: A submission is the complete set of files that are required in order to archive (aka ingest) the data. A submission includes the Readme.xml file, and all FITS and non-FITS files from the test that are intended to be archived. Think of “submission” as an envelope and its contents, where the contents are all the files being sent to the archive.

SubmissionType: SubmissionType is either PACKAGE or FILE. The archive uses the SubmissionType value to determine how to store and catalog the data. Select the option consistent with how the data will be used. When **SubmissionType = PACKAGE**, all the files in the submission will be tar’ed by the SID archive and ingested as a single entity. The only searchable information in the archive catalog will be the keywords taken from the Readme.xml file. Retrieval of the data will result in all the files (i.e., the entire tarball) being delivered. When **SubmissionType = FILE**, the submission must contain only FITS files. Each FITS file will be archived and cataloged. The searchable information for these FITS files will be the values for a very small subset of the FITS header keywords as well as the keyword values taken from the Readme.xml file. Each FITS file is

individually retrievable. That is, the entire set does not have to be retrieved in order to obtain one of the FITS files. Please note that non-FITS files may not be archived with SubmissionType = FILE.

Archive Catalog: The archive catalog is the database table(s) containing the metadata (i.e., the keyword values) ingested from the Readme.xml file and/or the FITS headers. It is the archive catalog that a user is browsing when using the Multi-Mission Archive at Space Telescope (MAST) search interface to locate data to retrieve. Note: the contents of the archive catalog are public, so anonymous searches are allowed.

DADS: DADS is the Data Archive and Distribution System. For the SID archive, DADS runs on a Linux system. The archive catalog is on a MS SQL Server database. DADS responds to requests for ingest of data (i.e., submissions) and for distribution (i.e., retrieval) of data. Both types of requests are generated for the user. The ingest request is generated by Operations when a submission is delivered. The distribution request is generated for the user via the Retrieval Options form, which is reached through the MAST JWST SID Archive (“SID Search”) interface.

Preparing Data for Archiving

Before submitting the data to the archive, determine the appropriate submission type, create the Readme.xml file, verify any FITS files in the submission, ensure the data are in the correct directory or directory structure, and the Readme.xml file is with the data in the top level directory. Set the permissions to allow the archive to copy the data from your directory to the archive directory.

Determine the SubmissionType

As noted in the Nomenclature section, there are 2 types of submissions, PACKAGE and FILE. There are advantages and disadvantages with each SubmissionType. Consider the future use of the test data as part of selecting a submission type.

A **PACKAGE** (SubmissionType) may contain many types of files, including FITS files. It is not limited to one type of file and the files may be contained in a hierarchical directory structure. The only searchable information available in the archive catalog is that taken from information in the Readme.xml file. The files are tar'd by the archive and archived as a tarball. When retrieving the data, the entire package (i.e., the tarball) will be distributed. Individual files cannot be retrieved from packages.

A **FILE** (SubmissionType) may contain only FITS files. These files must all be in the same directory with the Readme.xml file. The searchable information for the file submission includes the fields from the Readme.xml file and the required header keywords from the FITS files. Individual files may be retrieved from the archive. Any non-FITS files in the directory, except for the Readme.xml file, will cause ingest to fail, as will any non-valid FITS files.

The types of files produced during the test should determine the submission type to use when preparing the data for the archive.

- If there are no FITS files, use a SubmissionType of PACKAGE.
- If there are only FITS files, use a SubmissionType of FILE.

- If there are both FITS and non-FITS files, use both submission types – one for the FITS files and one for the other files.

While it is possible to include FITS files in a PACKAGE (SubmissionType), teams are encouraged to submit FITS files in a FILE (SubmissionType). This allows more flexibility in searching as additional keyword values are cataloged during ingest of the data.

When two submissions are planned for the same test, one for PACKAGE and one for FILE, use two Readme.xml files. Put each Readme.xml in a separate directory or directory structure with the corresponding data. In this case, the Readme.xml files should be identical except for the SubmissionType and Test Description. In the Test Description, state the data were archived in two different submissions. Provide a cross reference in each Test Description to the other submission.

Filename length limit

The archive has a hard limit on the filename length. It translates into a limit of 114 characters on the filename of any file that is sent to the archive. The archive will fail any request that contains filenames longer than this limit.

Create the Readme.xml file

All submissions for the archive must contain a Readme.xml file. The Readme.xml file may be generated with the help of a web-based tool or manually with the text editor of your choice. The teams are encouraged to use the web tool as it ensures the correctness of the xml and does some checking on the input values. Upon receipt, the archive will check that the Readme.xml file is valid xml. If it is not, the submission will be rejected. Users planning to manually create the Readme.xml file should be aware that both element names and enumerated elements are case sensitive. Consult the Readme.xml schema in Appendix A. For more information on XML conventions see <http://xml.silmaril.ie/authors/case/> .

Note: The Readme.xml file **must** be in the top level directory.

The web tool is currently available at <http://masthla.stsci.edu/jwst/> . At a future time the link will be changed to <http://mastjwst.stsci.edu/> . If neither of these links works, contact the archive help desk at sidarchive@stsci.edu . The figure below shows the opening page of the Readme file generator.

JWST I&T Data Archive Readme Generator

Area	<input type="text" value="ISIM Testbeds"/>	Facility where test is being conducted. If "Other", explain in Description.
Phase	<input type="text" value="Ambient"/>	Test conducted in ambient, cryo-vac, or other. If "Other", explain in Description.
Test Title	<input type="text"/>	As specified in the test plan
Test Number	<input type="text"/>	As specified in the test plan
Responsible Organization	<input type="text" value="NASA - JWST Project"/>	Organization conducting the test
Responsible Engineer	<input type="text"/>	Point of Contact for this test
Submission Type	<input type="text" value="Package"/>	'Package' or 'File' based units for the submission of test data to the SID archive.
Start Date	<input type="text"/> - <input type="text"/> - <input type="text"/>	Date and time in GMT at start of test (YYYY-MM-DD)
Start Time	<input type="text"/> : <input type="text"/> : <input type="text"/>	Date and time in GMT at start of test (HH:MM:SS)
End Date	<input type="text"/> - <input type="text"/> - <input type="text"/>	Date and time in GMT at end of test (YYYY-MM-DD)
End Time	<input type="text"/> : <input type="text"/> : <input type="text"/>	Date and time in GMT at start of test (HH:MM:SS)
Test Description	<input type="text"/>	Short text description of purpose/content of this test

Generator Version: 1.0.0.7
Deployed: 6/19/2013

Figure 1: Readme.xml file generator (<http://masthlabdev.stsci.edu/jwst/>)

Use the pull down menus, where available, to populate the fields. Otherwise fill in the requested information. If either Area or Phase is “Other,” include an explanation in the Test Description. Remember, it is the information in the Readme.xml file that will be searchable in the archive catalog. For a submission type of PACKAGE, this will be the *only* searchable information.

The information entered in the Test Title and Test Number fields will be used, along with the Area value to form the dataset name. There is a limit of 114 characters for the dataset name. Title is limited to 80 characters, Test Number to 20 characters and Area to 12 characters. There are two underscores in the dataset name.

Table 1 gives the list of Area values currently allowed by the Readme generator and their equivalent in the Readme.xml file. The Readme.xml column is important for users who do not use the Readme generator to

produce the xml file. PR 67766, SID 2011_2. At a future date, NIRISS will be added as an Organization and Area. (PR 77709)

Table 1 Example Area Values

In Readme Generator	In Readme.xml
Other	OTHER
ISIM Testbeds	ISIMTESTBEDS
ISIM I&T	ISIM_I_T
Spacecraft Testbeds	SCTESTBEDS
Spacecraft I&T	SC_I_T
Observatory I&T	OBS_I_T
Optical Telescope Element I&T	OTE_I_T
Operations	OPS
MIRI I&T	MIRI_I_T
NIRCam I&T	NIRCAM_I_T
NIRSpec I&T	NIRSPEC_I_T
TFI I&T	TFI_I_T
FGS I&T	FGS_I_T

Responsible Organization must be one of the values listed in Table 2. If the Readme.xml file does not contain a valid value for Responsible Organization, the submission will be rejected. **Note:** Group access to the data is based on the value of Responsible Organization. The Readme.xml column is important for users who do not use the Readme generator to produce the xml file.

Table 2 Allowed Responsible Organization Values

In Readme Generator	In Readme.xml
NASA-JWST Project	NASA
Northrup Grumman Aerospace Systems	NGAS
Space Telescope Science Institute	STSCI
MIRI SI Team	MIRI
NIRCAM SI Team	NIRCAM
NIRSPEC SI Team	NIRSPEC
TFI SI Team	TFI
FGS SI Team	FGS

The Start Date, Start Time, End Date, End Time, should be in GMT. If no times are entered by the user, the Readme Generator will error out, with messages indicating the problem written to the user’s screen and no Readme.xml file produced. The Readme Generator will check Start and End dates for existence, validity of dates, and sanity check for times in the future. Dates are **not** sanity checked for being too far in the past.

The Readme.xml file should be prepared by the Test conductor, the Designated Science Tester or some other responsible person designated by the team. The Readme file should be as accurate as possible, especially for packages. Think about how to describe the data in the package before creating the Readme.xml file. Use the Test Description field to identify the package/files uniquely. For example, “bias data at voltage = x, temperature = y, detectors <nnnn>.”

The Readme.xml file must be placed in the top level directory of the directory structure that contains the data.

Figure 2 shows an example Readme.xml file as produced by the Readme.xml file generator.

```
<?xml version="1.0" encoding="utf-8" ?>
<TestData xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="http://masthla.stsci.edu/jwst/readme.xsd">
  <Area>NIRCAM_I_T</Area>
  <Phase>CRYO</Phase>
  <StartDate>2010-08-20T17:42:02</StartDate>
  <EndDate>2010-08-20T17:42:02</EndDate>
  <Title>NIRCAM Image</Title>
  <Number>AD-24.0</Number>
  <Organization>NIRCAM</Organization>
  <Engineer>Richie Rich</Engineer>
  <Description>This is a test of the Readme generator.</Description>
  <SubmissionType>PACKAGE</SubmissionType>
</TestData>
```

Figure 2: Example Readme.xml File

Helpful hints for manual creation of Readme.xml file

Submitters of data should use the JWST Readme Generator web tool (see Figure 1) to create their Readme.xml files. Those choosing not to use the tool should exercise great care when creating the Readme.xml file manually, or editing an existing, valid, Readme.xml file, as the element names and values are case sensitive. Such users should also run an XML validator on their Readme.xml file to ensure the XML in this manually created or edited file is correct.

As noted above, there are restrictions on the values of Area, Phase and Responsible Organization. Tables 1 and 2 list the allowed values for Area and Responsible Organization. For Phase, the allowed values are CRYO, AMBIENT and OTHER. Submission type must be either PACKAGE or FILE. The times must be given in GMT. Values are required for Test Title, Test Number and Responsible Engineer. The Test Description should describe the data in sufficient detail so that there are no questions about the source of the data.

The xml schema for the Readme.xml file is listed in Appendix A. Please note the case sensitive nature of the element names and their enumerated values.

Verify the FITS files

A FITS verifier should be run on the FITS files in the FILE (SubmissionType) before the data are delivered to the archive. We encourage users to verify all FITS data, including those in Package submissions. This step is an important part of data preparation. HEASARC provides a FITS verifier named FITSverify, which is downloadable from their website at the following link.

<http://heasarc.gsfc.nasa.gov/docs/software/ftools/fitsverify/> . It can also be run online at

http://fits.gsfc.nasa.gov/fits_verify.html . The archive *will* run its version of FITSverify on the FITS files early

in the data ingest process. The submission will be rejected if a fits error is encountered for any of the required science header keywords (see Table 3) or if the fits error would prevent the extraction of the string value for any keyword in the header.

Note: FITS files in packages with SubmissionType = PACKAGE will **not** be verified for correctness by the archive.

The table below lists the required header keywords for the FITS files. If the FITS files were originally created by the FITSWriter tool, the values for these keywords (with the exception of TARGNAME) will already exist in the FITS file and will have appropriate values derived from the science telemetry. The keyword TARGNAME requires a user supplied value that should be provided to the FITSWriter at runtime or set by the user during later pre-ingest processing.

If the DATE keyword value is 00/00/00, the value will be replaced by a NULL in the database

Table 3 Required Science Header Keywords

Keyword	Description	Allowed Values or Example
DATE	date FITS file was created in UTC	Character string in date-time format yyyy-mm-ddThh:mm:ss
ROOTNAME	rootname of the observation set	FGSFITSTEST003071 (several files may have same root name)
INSTRUME	identifier for instrument used to acquire data	Character string e.g., NIRCAM, TFI
OBS_ID	Observation ID from the science data packet image header	Character string, taken from the science data package image header V12348001001P0000000000000
DATAMODE	Reference number used to specify ground data processing requirements for these data	Integer PR 67767
TARGNAME	Name of target	User supplied character string (alphanumeric)
EXPTIME	Observation exposure time (sss.cccc)	Floating point value, in seconds

Defaults for Required Science Header Keywords:

These keywords **MUST** appear in the primary header of each fits file targeted for the SID Archive. Blank or invalid entries for these required keywords will default to the values listed here.

The default values to be used by the SID Archive for the seven required keywords are as follows:

KEYWORD	DATA TYPE	DEFAULT VALUE
-----	-----	-----
DATE	datetime	Null
ROOTNAME	string	UNKNOWN
INSTRUME	string	UNKNOWN
OBS_ID	string	UNKNOWN
DATAMODE	integer	-1
TARGNAME	string	UNKNOWN
EXPTIME	float	Null

DATAMODE is constrained to the following values:

SI	Type	RANGE
MIRI	Flight, archived at STScI	0 - 24
NIRCam	Flight, archived at STScI	25 - 49
NIRSpec	Flight, archived at STScI	50 - 74
TFI/Guider	Flight, archived at STScI	75 - 99
MIRI	Ground, archived at STScI	65400 - 65424
NIRCam	Ground, archived at STScI	65425 - 65449
NIRSpec	Ground, archived at STScI	65450 - 65474
TFI/Guider	Ground, archived at STScI	65475 - 65499
Any	Ground, not archived at STScI	65500 - 65535

ROOTNAME must be present in the FITS headers. If the value of ROOTNAME is left blank, that is

- ROOTNAME= ,
- ROOTNAME=' ' or
- ROOTNAME=""

the value in the archive catalog will default to 'UNKNOWN.'

There is no FITS ROOTNAME for packages.

Delivering Data to the Archive

After the submission is prepared (Readme.xml generated and in the top level directory, data in the required directory structure, FITS files verified) it should be delivered to the SID archive for ingest. Click the "Contacts" link at the bottom of any MAST page, or go directly to <http://archive.stsci.edu/contacts.html> for directions on submitting data to the SID archive. Required information includes the location on the provider's system/disk of the data to be ingested.

Upon receipt of the data, the xml schema is validated, the FITS files in a file submission are verified and an ingest request is generated. When ingest completes, the team contact will be notified, by e-mail, of the success or failure of the submission.

Note: If the schema fails validation, the submission is rejected and the FITS files are not verified. FITS verification errors involving the required header keywords or extraction of the string value of any header keyword, will result in rejection of the submission. The team contact will be notified, by e-mail, of any such failures. The e-mail will indicate the problem files.

Dataset Naming Conventions in the SID Archive:

The teams should be aware that DADS has internal requirements on file uniqueness, which are handled through the dataset name (the `ippssoot` in HST parlance). One advantage of these requirements is that an updated version of a file may be archived over the previous version. Both versions remain in the archive, but only the most recent, the so-called “best version,” is visible via the MAST interface. These internal requirements also place limits on the length of the Test Title and Test Number. Title is limited to 80 characters, Test Number to 20 characters and AREA to 12 characters.

For both PACKAGE and File submission types, the dataset name syntax is Area-TestTitle-TestNumber. Previous versions of PACKAGES and FITS files may be superseded by delivering the improved version with a `Readme.xml` that duplicates the Area, Title and Number values in the original `Readme` file and, in the FITS file, a `ROOTNAME` value identical to that in the original FITS file.

Note: Before redelivering any files or packages, contact the SID archive at sidarchive@stsci.edu .

SID Archive ITAR Handling:

The SID Archive will not contain ITAR information. Test submissions containing ITAR related material require JWST project office approval prior to delivery of the submission to the SID Archive. For more detailed information on SID Archive submissions related to ITAR, contact sidarchive@stsci.edu .

Searching for Data in the SID Archive

Access to the SID archive is through the MAST interface <http://archive.stsci.edu/jwst/sidarchive/search.php> . Clicking on this link brings up a standard MAST web page, as shown in Figure 3.

The MAST SID Search web page has two basic parts. The upper portion of the page allows the user to enter specific details about the data being sought. Values may be entered for, among others, Instrument, Test Number and Data Mode. The type of Submission, PACKAGE or FILE, can also be indicated. In addition, there are six User-specified fields, which allow the user to add, as a form element, any column from the archive catalog. For more information about each field click on the link "Field Descriptions". This brings up a page, titled “SID Field Descriptions,” of all the available columns and gives the database column name, the column label, the data type, example values or a list of valid values in that field, and the data type. The page usually displays in a separate browser tab.

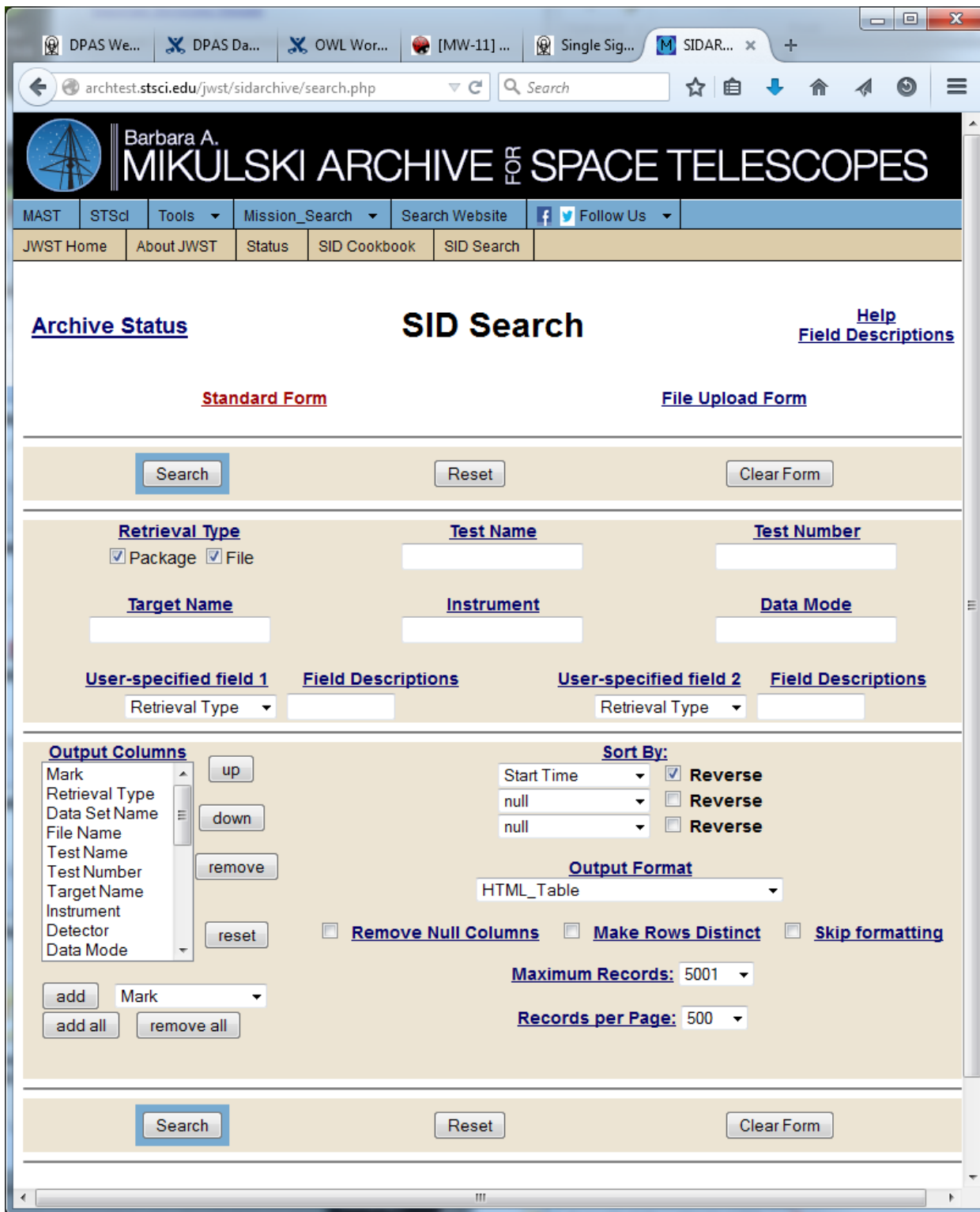


Figure 3: Standard MAST Web Page

The lower section of the web page, also known as the output options section, allows the user to add, remove and/or reorder the fields to be displayed in the output, and to format the output of the search results. On the left side is a list of columns to display with the search results. Add columns by selecting them from the pull down menu below the “Output Columns” box. Remove any of the columns by highlighting that column and clicking on the remove button to the right. Change the order of the columns by clicking on field in the output columns

box and click on the up and down buttons to the right. Clicking on the reset button will restore the default output columns settings. Output may be sorted by up to three columns. Selectable output formats include HTML, comma separated value text, Excel spread sheet, and VO table format. The HTML format is the format that will give access to data retrieval options.

For users who have a prepared list of Data Set Names, Test Names or Test Numbers to search, the File Upload Form allows input of such files. Click the File Upload Form link in the upper right of the web page. Use the top section of the form to describe the format of a file containing a list of Data Set Names or Test Names or Test Numbers. Click on the column labels for additional help. The local file name is the name of the file on the user's system. The file must be an ASCII text file or table with one entry per line with the fields (e.g. Test Title or Test Name) separated with one of the allowed delimiters. Use the form to designate in which column the Test Title or Test Name is found. If appropriate, designate the delimiter (tab, |, comma (,), semi-colon(;)).

After specifying the search criteria, click on the Search button to initiate the search.

Retrieving Data

While anyone may search the SID archive, only STScI SSO credentialed and authorized users may retrieve data. At the time this document was written, it is expected that all STScI employees and all members of a Responsible Organization will have their accounts converted to STScI SSO credentials. You will receive e-mail if your account is converted to an STScI SSO account. See STScI's Single Sign On (SSO) Portal, <https://ssoportal.stsci.edu/cas/login>, for details on recovering a forgotten password, updating your SSO account information (e.g., contact information, change password, etc.) and other useful information.

Note: If the user is already listed as a member of a Responsible Organization (see Table 2 for a list of the Responsible Organizations) the value in the "username" field must match the corresponding value provided by the Responsible Organization. To be an authorized user, the user must be a member of one of the Responsible Organizations. If the user is not a member of a Responsible Organization, the user should have the head of the Responsible Organization contact the archive help desk, sidarchive@stsci.edu, to request addition of the user to the Responsible Organization list. Again, the "username" must be the same on the STScI SSO account and in the Responsible Organization list.

Figure 4 shows the MAST web page set up to search for files and packages related to test A_SBF_SW_f20_pupil_109f_0006. To see all submissions, both packages and files, both the "package" and "file" box under "Retrieval Type," must be checked. While the example includes the complete test title, a wild card (% or *) could also be used, say A_SBF_SW_f20_pupil_109f%.

Clicking the Search button initiates the search. When the search is complete, the results are displayed on the Search Results Page, as shown in Figure 5. If the user requested CSV, VOTable or excel spreadsheet output, the results would be written to the user's disk in the appropriate file type. The listed results include both files (File Name extension .fits) and packages (file name extension .tar).

Note the "Mark" check box. Indicate the data to retrieve by clicking the appropriate "Mark" check box. To select all the results, use the "Mark all" button. Submit the request for validation by clicking the "Submit Marked data for retrieval" button.

[Archive Status](#) **SID Search** [Help](#)
[Field Descriptions](#)

[Standard Form](#) [File Upload Form](#)

Retrieval Type <input checked="" type="checkbox"/> Package <input checked="" type="checkbox"/> File	Test Name SW_f20_pupil_109f_0006	Test Number <input type="text"/>
Target Name <input type="text"/>	Instrument <input type="text"/>	Data Mode <input type="text"/>
User-specified field 1 Retrieval Type <input type="text"/>	Field Descriptions <input type="text"/>	User-specified field 2 Retrieval Type <input type="text"/>
Field Descriptions <input type="text"/>	Field Descriptions <input type="text"/>	Field Descriptions <input type="text"/>

Output Columns	Sort By:	
<input type="button" value="up"/> <input type="button" value="down"/> <input type="button" value="remove"/> <input type="button" value="reset"/>	Start Time <input type="text"/> <input checked="" type="checkbox"/> Reverse null <input type="text"/> <input type="checkbox"/> Reverse null <input type="text"/> <input type="checkbox"/> Reverse	
<input type="button" value="add"/> Mark <input type="text"/> <input type="button" value="add all"/> <input type="button" value="remove all"/>	Output Format	
	HTML_Table <input type="text"/>	
	<input type="checkbox"/> Remove Null Columns	<input type="checkbox"/> Make Rows Distinct
		<input type="checkbox"/> Skip formatting
	Maximum Records: 5001 <input type="text"/>	
	Records per Page: 500 <input type="text"/>	

Figure 4: Standard MAST Web Page set up to search for test A_SBF_SW_f20_pupil_109f_0006



JWST SID Search Results

Edit Query

/OPlot is currently unavailable.

number of rows returned = 4

note: reload page if no results are shown

Submit marked data for retrieval

Mark all Unmark all Mark public Unmark public Mark proprietary Unmark proprietary

Mark	Instrument	Data Set Name	Test Number	Test Name	Date
<input type="checkbox"/> @	OMA	NIRCAM_I_T-A_SBF_SW_F20_PUPIL_109F_0006-NRCFM1_OMA100046	nrcfm1_oma100046	A_SBF_SW_f20_pupil_109f_0006	2013-06-07 15:50:32
<input type="checkbox"/> @	OMA	NIRCAM_I_T-A_SBF_SW_F20_PUPIL_109F_0006-NRCFM1_OMA100036	nrcfm1_oma100036	A_SBF_SW_f20_pupil_109f_0006	2013-06-07 15:50:20
<input type="checkbox"/> @		NIRCAM_I_T-A_SBF_SW_F20_PUPIL_109F_0006-NRCFM1_OMA100046-V1	nrcfm1_oma100046	A_SBF_SW_f20_pupil_109f_0006	
<input type="checkbox"/> @		NIRCAM_I_T-A_SBF_SW_F20_PUPIL_109F_0006-NRCFM1_OMA100036-V1	nrcfm1_oma100036	A_SBF_SW_f20_pupil_109f_0006	

Submit marked data for retrieval

Mark all Unmark all Mark public Unmark public Mark proprietary Unmark proprietary

Figure 5: The Search Results Page

A second page will be displayed, as shown in Figure 6. Note the yellow highlighting on the “Mark” check box. This indicates that the data are proprietary. They may only be retrieved by someone who is authorized to retrieve proprietary data for this instrument. See below for more information on proprietary data in the SID archive.

Datasets Marked for Retrieval

[Archive Status](#)

NEW [Important Downtime Message](#) **NEW**

4 datasets (0 public, 4 proprietary) marked.

Submit marked data for retrieval from DADS

Mark all Unmark all Mark proprietary Unmark proprietary Mark public Unmark public

Row	Mark	Rootname	Filename	Release	Row
1	<input checked="" type="checkbox"/> @	NIRCAM_I_T-A_SBF_SW_F20_PUPIL_109F_0006-NRCFM1_OMA100036	A_SBF_SW_f20_pupil_109f_0006_20120623_191104.fits	2038-01-01 01:01:01	1
2	<input checked="" type="checkbox"/> @	NIRCAM_I_T-A_SBF_SW_F20_PUPIL_109F_0006-NRCFM1_OMA100036-V1	nircam_i_t-a_sbf_sw_f20_pupil_109f_0006-nrcfm1_oma100036-v1.tar	2038-01-01 01:01:01	2
3	<input checked="" type="checkbox"/> @	NIRCAM_I_T-A_SBF_SW_F20_PUPIL_109F_0006-NRCFM1_OMA100046	A_SBF_SW_f20_pupil_109f_0006_20120623_191434.fits	2038-01-01 01:01:01	3
4	<input checked="" type="checkbox"/> @	NIRCAM_I_T-A_SBF_SW_F20_PUPIL_109F_0006-NRCFM1_OMA100046-V1	nircam_i_t-a_sbf_sw_f20_pupil_109f_0006-nrcfm1_oma100046-v1.tar	2038-01-01 01:01:01	4

Submit marked data for retrieval from DADS

Mark all Unmark all Mark proprietary Unmark proprietary Mark public Unmark public

Thu Jun 5 12:35:22 2014
archive@stsci.edu

Copyright © 2014 AURA, Inc.
All Rights Reserved.

Figure 6: Datasets to Retrieve

Click the “Submit marked data for retrieval from DADS” button. The Retrieval options page is displayed. On this page the user indicates who they are, how the data should be delivered, and, if they are to be sftp’ed, the location and account information for the sftp transfer. An example is shown in Figure 7.

Since all data in the SID archive are proprietary, the user must identify themselves via their STScI SSO credentials (i.e, Archive Username must be your STScI SSO Account name and Archive Password must be the password for your STScI SSO Account).

To send the retrieval request to DADS, click on the “Send retrieval request to ST-DADS” button. The final page displayed is the Request Sent to ST-DADS page, as shown in Figure 8. If there is an error on the Retrieval Options page, say an incorrect password or account information, an error message will be displayed instead of the Request Sent page. The user may use the back arrow on the browser to return to the Retrieval Options page to re-enter the required information.

Retrieval Options

[Archive Status](#)
[Important Downtime Message](#) 1 dataset (0 public, 1 proprietary) marked.

Archive Username STScI SSO Account	Archive Password ●●●●●●●●
Delivery options <input checked="" type="radio"/> FTP: FTP the data to the destination shown <input checked="" type="checkbox"/> Use sftp (OpenSSH v2) <input type="radio"/> STAGE: Put the data onto the Archive staging disk* <input type="radio"/> DVD: Send the data to me on DVD. <input type="checkbox"/> Compress the files using gzip.	Destination (if you selected FTP): Hostname Your machine Directory Destination directory Username Your machine account Password ●●●●●●●●
<input type="button" value="Send retrieval request to ST-DADS"/>	<input type="button" value="Reset form to default values"/>

Fri Jan 23 11:01:39 2015
archive@stsci.edu

Copyright © 2015 AURA, Inc.
All Rights Reserved.

Figure 7: The Retrieval Options Page



Figure 8: Request Accepted

The user will receive e-mail regarding the success or failure of the archive request. When data retrieval is complete, the user will receive an additional message giving the status of the retrieval. Figure 9 shows a typical message confirming receipt of the retrieval request, while Figure 10 shows the message for the retrieval response.

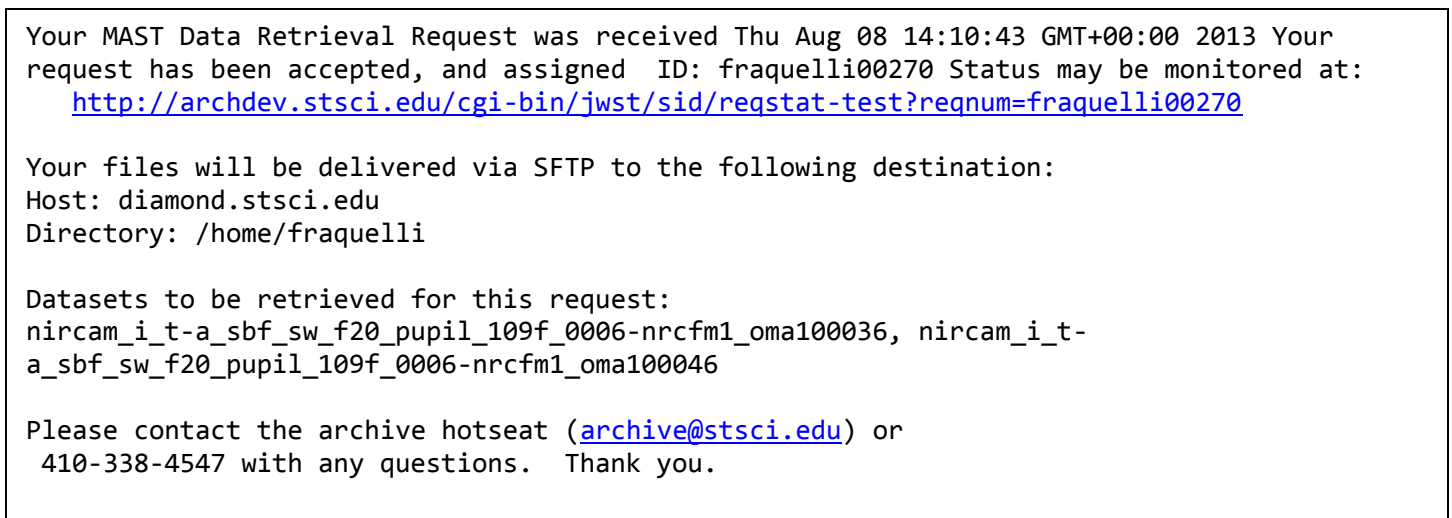


Figure 9: Retrieval Request Confirmation E-mail

Your request fraquelli00270
was completed at Thu Aug 08 14:10:55 GMT+00:00 2013 REQUEST DELIVERY STATUS: SUCCESS

Your data have been shipped to
[fraquelli@diamond.stsci.edu:/home/fraquelli](mailto:fraquelli@diamond.stsci.edu)

Total size of delivered files: 1.274 MB (1 Mb = 1024 kb)

Delivery status of individual files

A_SBF_SW_f20_pupil_109f_0006_20120623_191104.fits with state: success (0.637 Mb)
A_SBF_SW_f20_pupil_109f_0006_20120623_191434.fits with state: success (0.637 Mb)

Please contact the archive hotseat (archive@stsci.edu) or
410-338-4547 with any questions. Thank you.

Figure 10: Retrieval Response Confirmation E-mail

Proprietary Access

When a user requests proprietary data, a check is made to ensure that user is allowed access to the data. This means the user must supply their STScI SSO credentials on the Retrieval Options page (Figure 7). For the SID archive, proprietary rights are managed via group access, with “Responsible Organization” as the group (PR 64884). See Table 2 for a list of values for Responsible Organization. If the user is not recognized as a member of the Responsible Organization associated with the data, access to the data will be denied. This is to prevent, say, a MIRI user from obtaining NIRSPEC data.

Appendix A Readme.xml need to update for PR 77709 (NIRISS vs TFI)

The schema for the Readme.xml file is provided for the convenience of those who choose to create their own xml or who edit existing xml. Note the schema includes all valid values for Area, Phase and Responsible Organization. Users of this schema will need to remove the values that do not apply.

```
<?xml version="1.0" encoding="utf-8" ?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema" elementFormDefault="unqualified">
  <xsd:element name="TestData">
    <xsd:complexType>
      <xsd:sequence>
        <xsd:element name="Area" type="areaType"> </xsd:element>
        <xsd:element name="Phase" type="phaseType"> </xsd:element>
        <xsd:element name="StartDate" type="xsd:dateTime"> </xsd:element>
        <xsd:element name="EndDate" type="xsd:dateTime"> </xsd:element>
        <xsd:element name="Title" type="xsd:string"> </xsd:element>
        <xsd:element name="Number" type="xsd:string">
          <xsd:annotation>
            <xsd:documentation>
              This should be changed to the relevant number type, if the data
              does not contain any letters or period demarcators.
            </xsd:documentation>
          </xsd:annotation>
        </xsd:element>
        <xsd:element name="Organization" type="organizationType"> </xsd:element>
        <xsd:element name="Engineer" type="xsd:string"> </xsd:element>
        <xsd:element name="Description" type="xsd:string"> </xsd:element>
        <xsd:element name="SubmissionType" type="submissionType"> </xsd:element>
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
  <xsd:simpleType name="areaType">
    <xsd:restriction base="xsd:string">
      <xsd:enumeration value="ISIMTESTBEDS"> <!-- ISIM Testbeds --> </xsd:enumeration>
      <xsd:enumeration value="ISIM_I_T"> <!-- ISIM I&T --> </xsd:enumeration>
      <xsd:enumeration value="SCTESTBEDS"> <!-- Spacecraft Testbeds --> </xsd:enumeration>
      <xsd:enumeration value="SC_I_T"> <!-- Spacecraft I&T --> </xsd:enumeration>
      <xsd:enumeration value="OBS_I_T"> <!-- Observatory I&T --> </xsd:enumeration>
      <xsd:enumeration value="OTE_I_T"> <!-- Optical Telescope Element I&T --> </xsd:enumeration>
      <xsd:enumeration value="OPS"> <!-- Operations --> </xsd:enumeration>
      <xsd:enumeration value="MIRI_I_T"> <!-- MIRI I&T --> </xsd:enumeration>
      <xsd:enumeration value="NIRCAM_I_T"> <!-- NIRC&I I&T --> </xsd:enumeration>
      <xsd:enumeration value="NIRSPEC_I_T"> <!-- NIRSpec I&T --> </xsd:enumeration>
      <xsd:enumeration value="TFI_I_T"> <!-- TFI I&T --> </xsd:enumeration>
      <xsd:enumeration value="FGS_I_T"> <!-- FGS I&T --> </xsd:enumeration>
      <xsd:enumeration value="OTHER"> <!-- Information in Description field --> </xsd:enumeration>
    </xsd:restriction>
  </xsd:simpleType>
</xsd:schema>
```

```

</xsd:simpleType>
<xsd:simpleType name="phaseType">
  <xsd:restriction base="xsd:string">
    <xsd:enumeration value="AMBIENT"> <!-- Test conducted at ambient --> </xsd:enumeration>
    <xsd:enumeration value="CRYO"> <!-- Test conducted in cryo-vac --> </xsd:enumeration>
    <xsd:enumeration value="OTHER"> <!-- Information in Description field --> </xsd:enumeration>
  </xsd:restriction>
</xsd:simpleType>
<xsd:simpleType name="organizationType">
  <xsd:restriction base="xsd:string">
    <xsd:enumeration value="NASA"> <!-- JWST Project --> </xsd:enumeration>
    <xsd:enumeration value="NGAS"> <!-- Northrup Grumman Aerospace Systems --> </xsd:enumeration>
    <xsd:enumeration value="STSCI"> <!-- Space Telescope Science Institute --> </xsd:enumeration>
    <xsd:enumeration value="MIRI"> <!-- The MIRI SI Team --> </xsd:enumeration>
    <xsd:enumeration value="NIRCAM"> <!-- The NIRCcam SI Team --> </xsd:enumeration>
    <xsd:enumeration value="NIRSPEC"> <!-- The NIRSspec SI Team --> </xsd:enumeration>
    <xsd:enumeration value="TFI"> <!-- The TFI SI Team --> </xsd:enumeration>
    <xsd:enumeration value="FGS"> <!-- The FGS SI Team --> </xsd:enumeration>
  </xsd:restriction>
</xsd:simpleType>
<xsd:simpleType name="packageType">
  <xsd:restriction base="xsd:string">
    <xsd:enumeration value="PACKAGE"> <!-- Directory structure containing data to archive
                                     as a single package -->
    </xsd:enumeration>
    <xsd:enumeration value="FILE"> <!-- Individual FITS files --> </xsd:enumeration>
  </xsd:restriction>
</xsd:simpleType>
</xsd:schema>

```