STScI | SPACE TELESCOPE SCIENCE INSTITUTE

EXPANDING THE FRONTIERS OF SPACE ASTRONOMY

Reproducibility and Provenance of James Webb Space Telescope Data Products

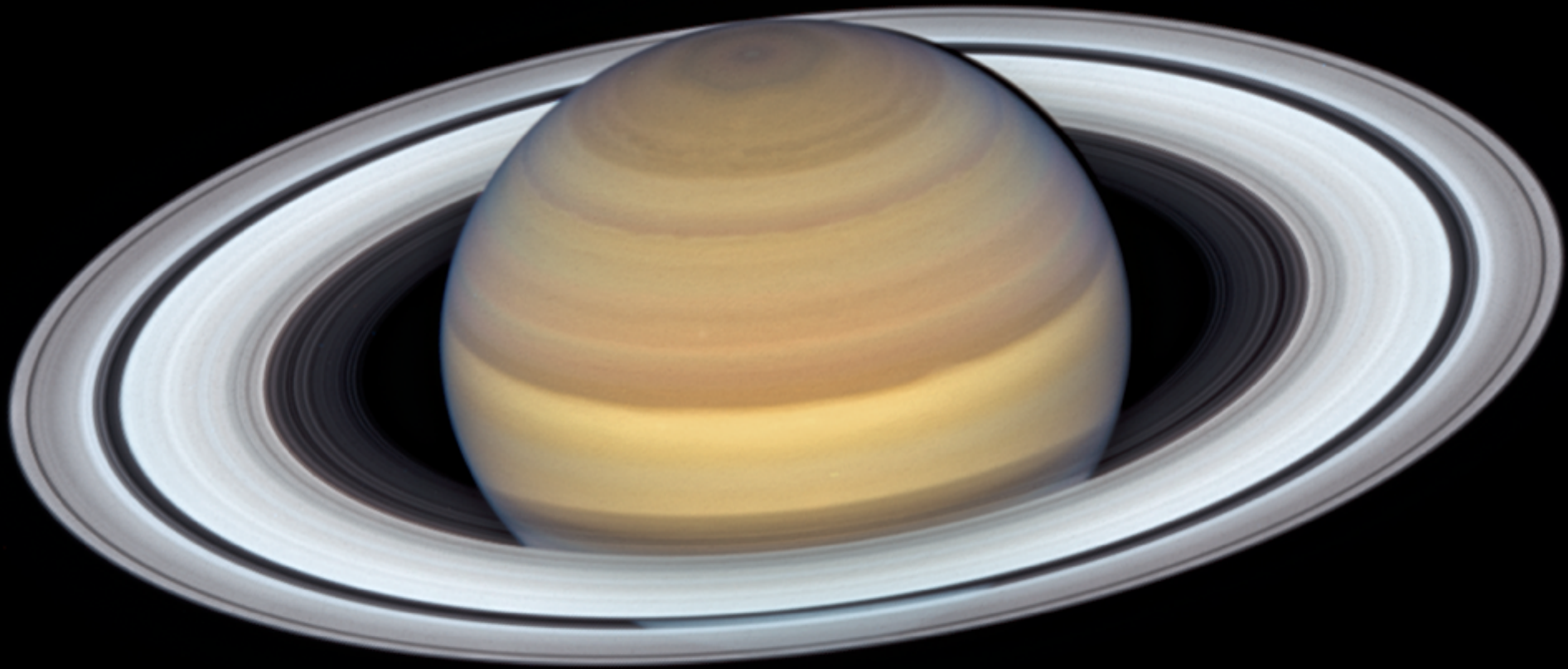
Howard Bushouse
Science Software Branch/STScI

SC19/CRE2019 17 Nov 2019


James Webb


STScI

The Science Images We’re Used to Seeing from Hubble



Stephan's Quintet

Nice eye candy from the Office of Public Outreach

But that's ***NOT*** what we get from the telescope

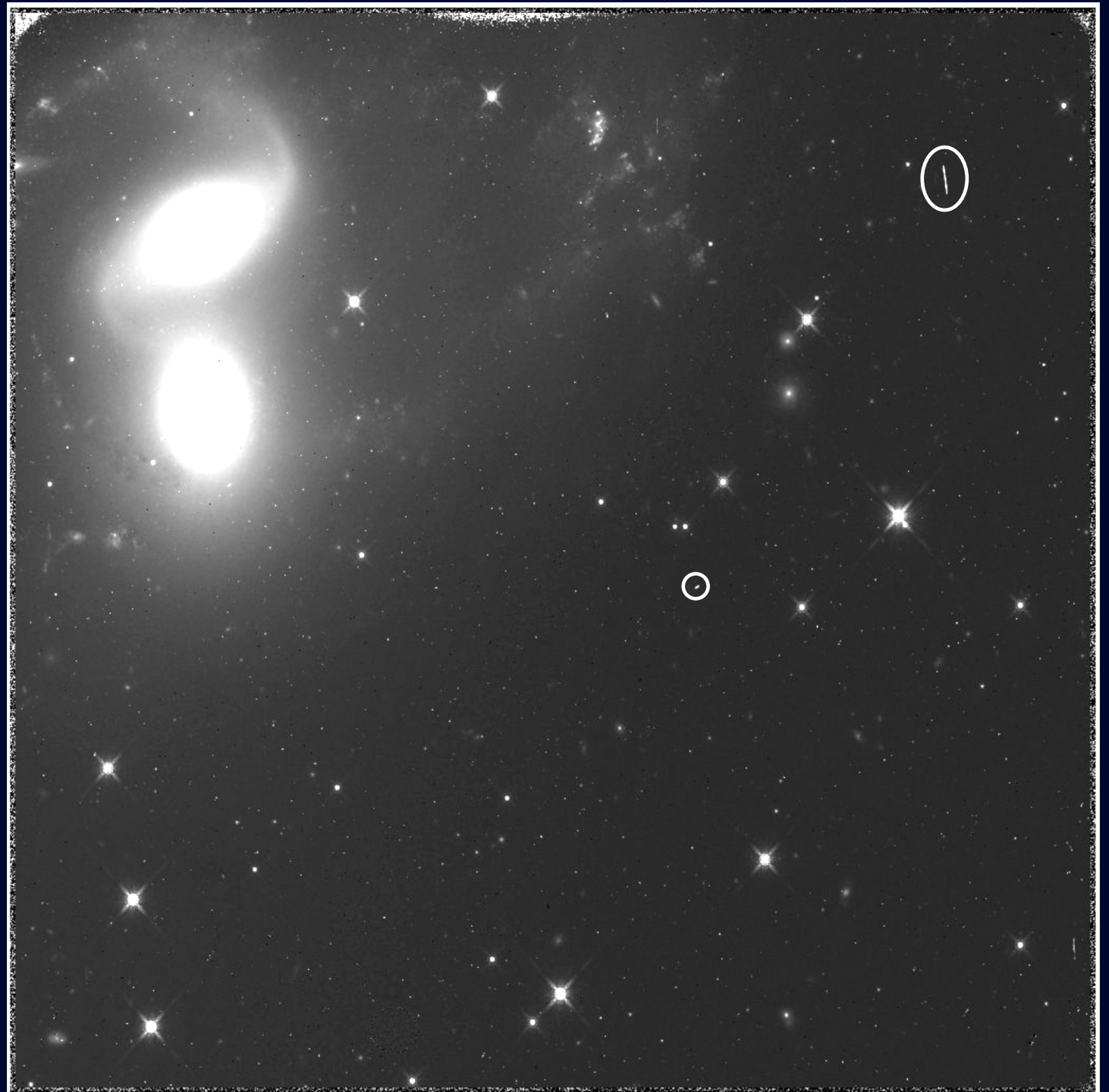
Stephan's Quintet

What we actually get from the telescope:

Raw infrared detector readout from the Hubble Space Telescope Wide-Field Camera 3

Stephan's Quintet

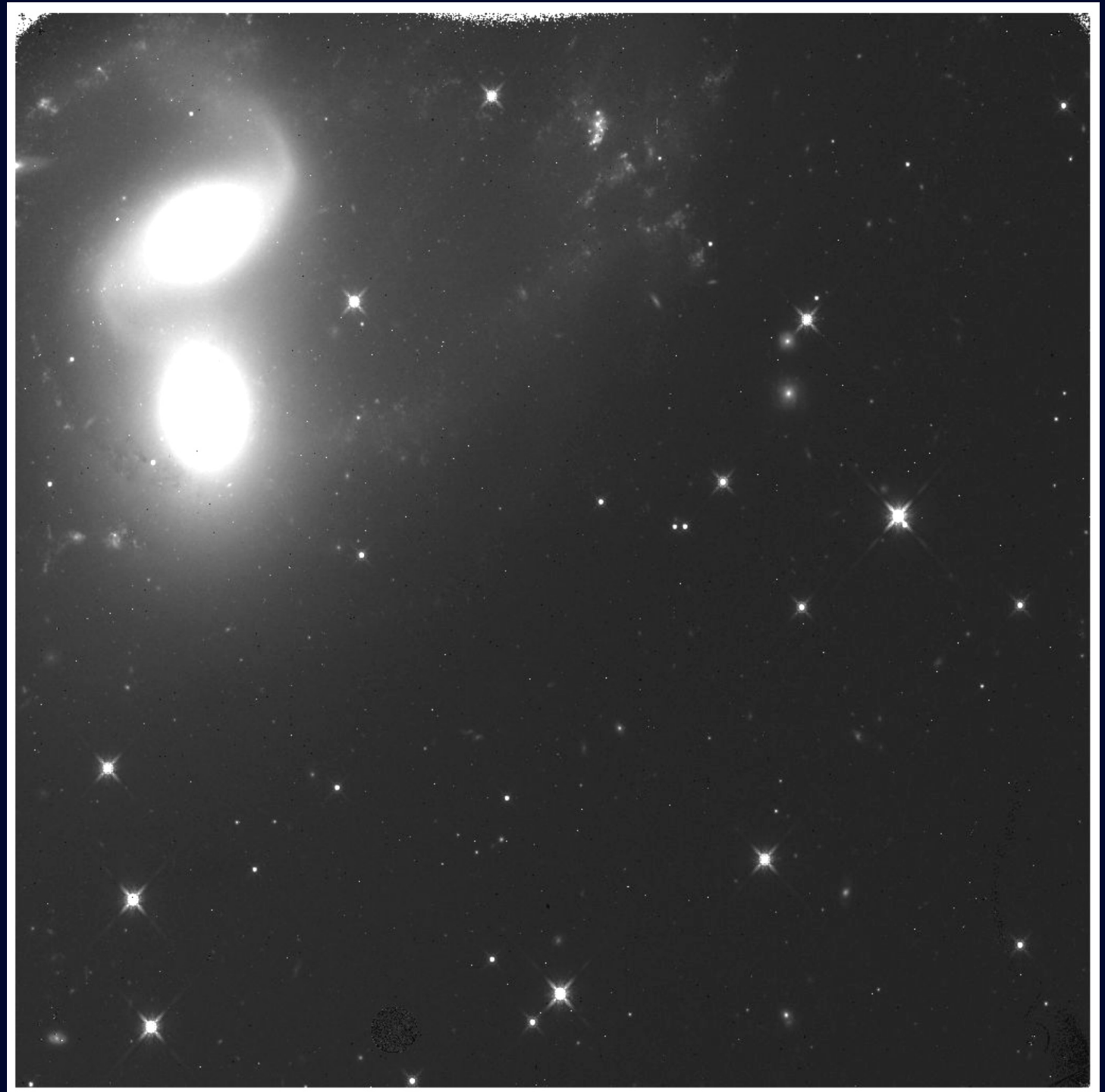
After removing detector and electronic features



Stephan's Quintet

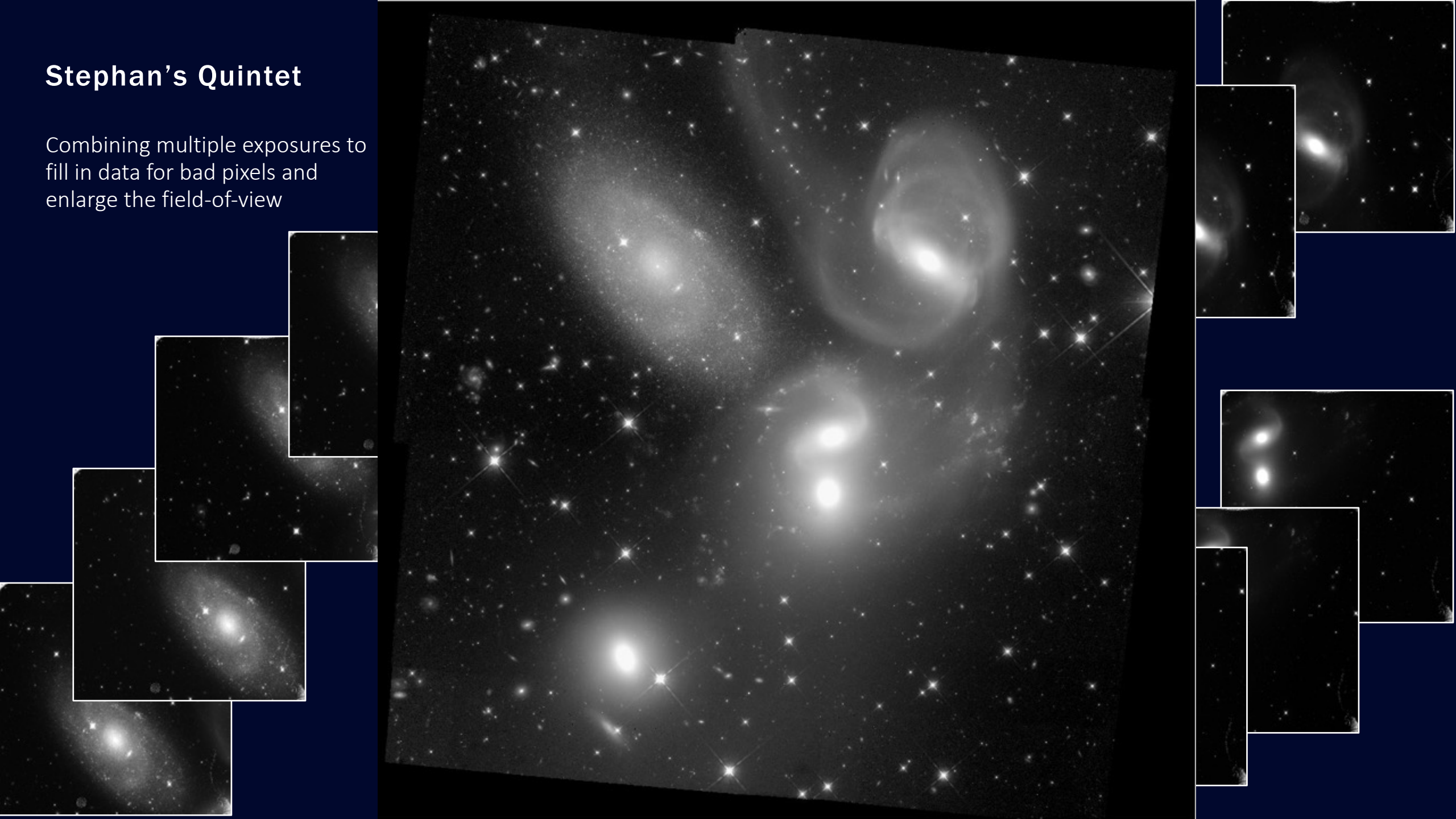
After removing Cosmic Ray hits*

(*No photons were harmed in the cleaning of this image)



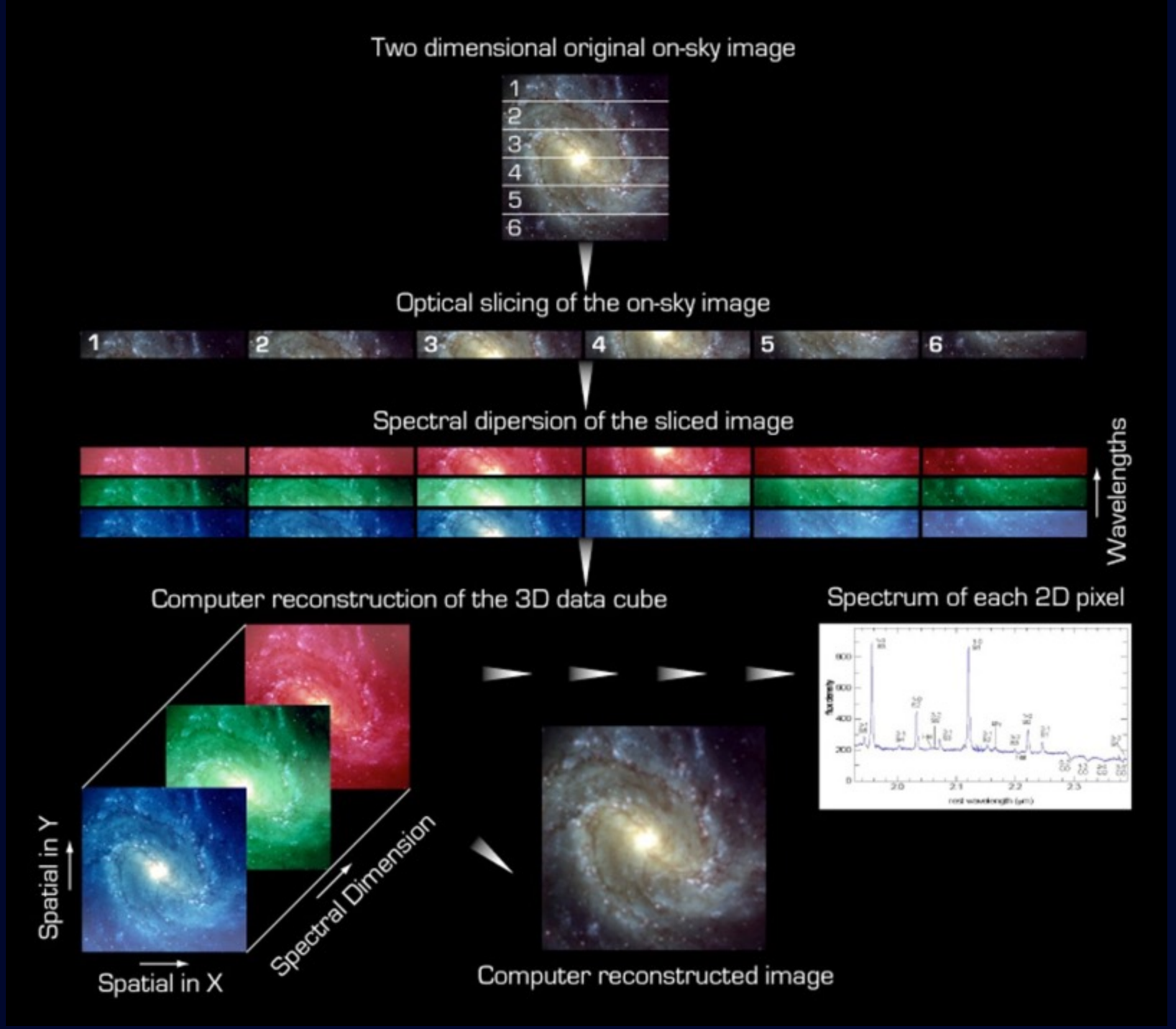
Stephan's Quintet

Combining multiple exposures to fill in data for bad pixels and enlarge the field-of-view



Integral Field Unit 3-D Spectroscopy

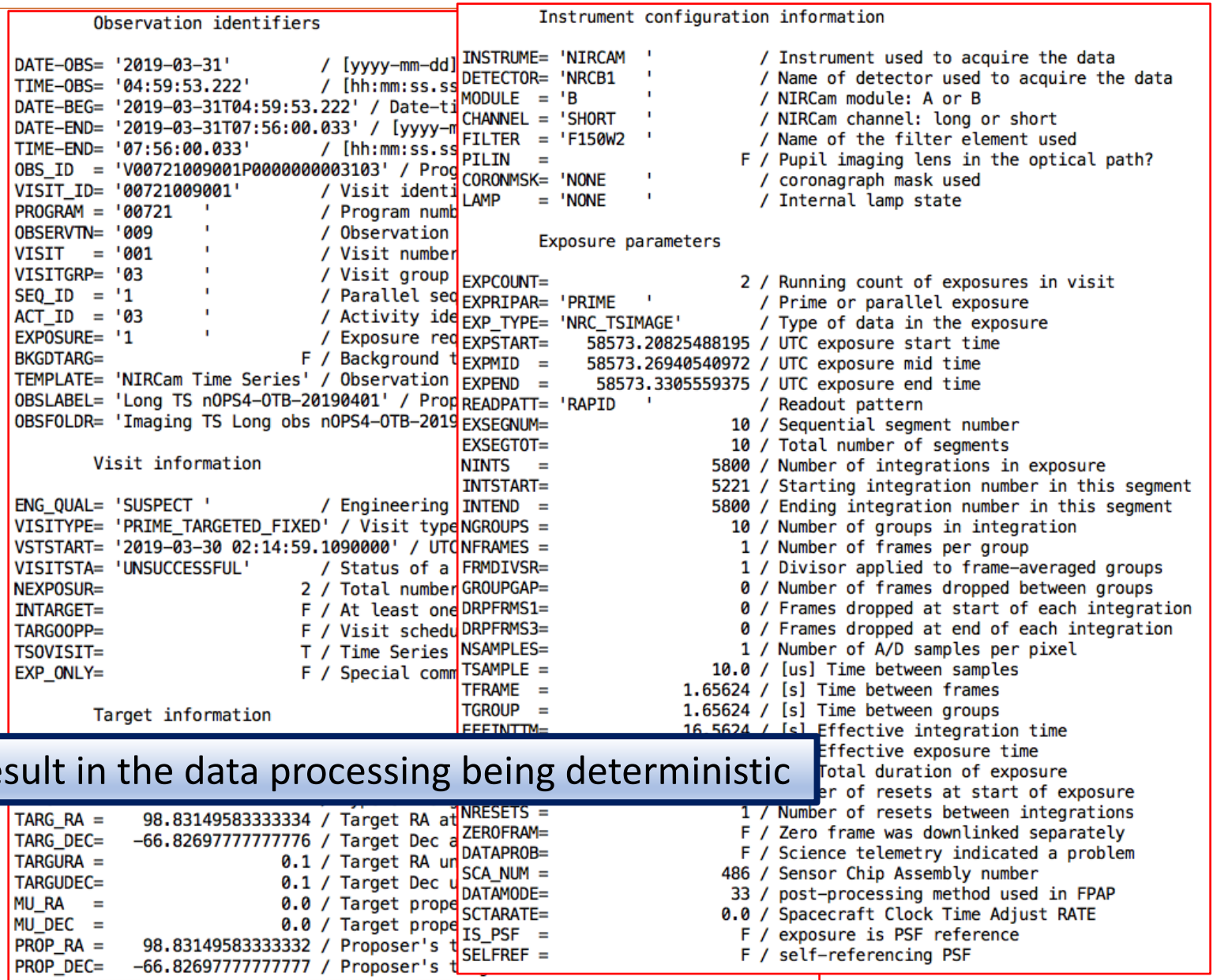
Sky field is (literally) sliced into pieces and each slice sent through a spectrograph




Data Processing Pipelines

Hubble and Webb Data Processing

- Exposure Telemetry
- Sorting & Repackaging
- Raw FITS* File:
  meta data
  image data (pixel values)
- Corrected/Calibrated Files
- Multi-Mission Archive



```
        Observation identifiers

DATE-OBS= '2019-03-31'         / [yyyy-mm-dd]
TIME-OBS= '04:59:53.222'       / [hh:mm:ss.ss
DATE-BEG= '2019-03-31T04:59:53.222' / Date-ti
DATE-END= '2019-03-31T07:56:00.033' / [yyyy-m
TIME-END= '07:56:00.033'       / [hh:mm:ss.ss
OBS_ID  = 'V00721009001P0000000003103' / Prog
VISIT_ID= '00721009001'        / Visit identi
PROGRAM = '00721   '           / Program numb
OBSERVTN= '009     '           / Observation
VISIT   = '001     '           / Visit number
VISITGRP= '03      '           / Visit group
SEQ_ID  = '1       '           / Parallel seq
ACT_ID  = '03      '           / Activity ide
EXPOSURE= '1       '           / Exposure req
BKGDTARG=                    F / Background t
TEMPLATE= 'NIRCam Time Series' / Observation
OBSLABEL= 'Long TS nOPS4-OTB-20190401' / Prop
OBSFOLDR= 'Imaging TS Long obs nOPS4-OTB-2019

        Visit information

ENG_QUAL= 'SUSPECT '           / Engineering
VISITYPE= 'PRIME_TARGETED_FIXED' / Visit type
VSTSTART= '2019-03-30 02:14:59.1090000' / UTC
VISITSTA= 'UNSUCCESSFUL'       / Status of a
NEXPOSUR=                    2 / Total number
INTARGET=                    F / At least one
TARGOOPP=                    F / Visit schedu
TSOVISIT=                    T / Time Series
EXP_ONLY=                    F / Special comm

        Target information

TARG_RA =    98.83149583333334 / Target RA at
TARG_DEC=   -66.82697777777776 / Target Dec a
TARGURA =                  0.1 / Target RA un
TARGUDEC=                  0.1 / Target Dec u
MU_RA   =                  0.0 / Target prope
MU_DEC  =                  0.0 / Target prope
PROP_RA =    98.83149583333332 / Proposer's t
PROP_DEC=   -66.82697777777777 / Proposer's t
```

```
        Instrument configuration information

INSTRUME= 'NIRCAM  '           / Instrument used to acquire the data
DETECTOR= 'NRCB1   '           / Name of detector used to acquire the data
MODULE  = 'B       '           / NIRCam module: A or B
CHANNEL = 'SHORT   '           / NIRCam channel: long or short
FILTER  = 'F150W2  '           / Name of the filter element used
PILIN   =                    F / Pupil imaging lens in the optical path?
CORONMSK= 'NONE    '           / coronagraph mask used
LAMP    = 'NONE    '           / Internal lamp state

        Exposure parameters

EXPCOUNT=                    2 / Running count of exposures in visit
EXPRIPAR= 'PRIME   '           / Prime or parallel exposure
EXP_TYPE= 'NRC_TSIMAGE'        / Type of data in the exposure
EXPSTART=    58573.20825488195 / UTC exposure start time
EXPMID  =    58573.26940540972 / UTC exposure mid time
EXPEND  =     58573.3305559375 / UTC exposure end time
READPATT= 'RAPID   '           / Readout pattern
EXSEGNUM=                   10 / Sequential segment number
EXSEGTOT=                   10 / Total number of segments
NINTS   =                 5800 / Number of integrations in exposure
INTSTART=                 5221 / Starting integration number in this segment
INTEND  =                 5800 / Ending integration number in this segment
NGROUPS =                   10 / Number of groups in integration
NFRAMES =                    1 / Number of frames per group
FRMDIVSR=                    1 / Divisor applied to frame-averaged groups
GROUPGAP=                    0 / Number of frames dropped between groups
DRPFRMS1=                    0 / Frames dropped at start of each integration
DRPFRMS3=                    0 / Frames dropped at end of each integration
NSAMPLES=                    1 / Number of A/D samples per pixel
TSAMPLE =                 10.0 / [us] Time between samples
TFRAME  =              1.65624 / [s] Time between frames
TGROUP  =              1.65624 / [s] Time between groups
EFFINTTM=             16.5624 / [s] Effective integration time
                                 Effective exposure time
                                 Total duration of exposure
                                 er of resets at start of exposure
NRESETS =                    1 / Number of resets between integrations
ZEROFRAM=                    F / Zero frame was downlinked separately
DATAPROB=                    F / Science telemetry indicated a problem
SCA_NUM =                  486 / Sensor Chip Assembly number
DATAMODE=                   33 / post-processing method used in FPAP
SCTARATE=                  0.0 / Spacecraft Clock Time Adjust RATE
IS_PSF  =                    F / exposure is PSF reference
SELFREF =                    F / self-referencing PSF
```

Static meta data result in the data processing being deterministic

*FITS = Flexible Image Transport System

Series of Individual Correction/Calibration Steps

Example Pipeline Steps	Calibration Reference Files
Flag Bad Pixels	Bad Pixel Mask
Correct for Electronic Bias Drift	Reference Pixel Parameters
Flag Saturated Pixels	Saturation Threshold Map
Detect and Flag Cosmic Ray Hits	Detector Noise Parameters
Attach Telescope Pointing Information	Science Instrument Aperture File
Absolute Flux Calibration	Flux Conversion Factors
Distortion Correction	Distortion Models

Reference files often change as a result of instrument/detector aging, as well as better understanding and analysis of calibration data.

Careful selection and tracking of reference files is therefore important.

Reference Data Access

JWST Calibration Reference Data System (CRDS)

Obtain Best Reference Files

1. Using the Command Line
2. From Dataset ID or FITS Header Upload
3. Exploring with Instrument Parameters

When necessary, pipeline software versions are tied to CRDS Context values

Operational References (under context jwst_0565.pmap)

- fgs
- miri
- nircam
- niriss
- nirspec
- system

Context History (more history, all contexts)

Start Date	Context	Status	Description
2019-11-08	jwst_0565.pmap	operational	Removes files from the NIRSpec PHOTOM rmap that needed to be replaced with updated reference files that were delivered on October 23, 2019. Sets the updated reference file USEAFTER to 2010-01-01, correcting dates which only covered recent data. See JIRA issue CRDS-310 for more information.
2019-11-01	jwst_0563.pmap	delivered	These new NIRCam distortion reference files were created using the astrometric solution contained in the PRD as of 24 Oct 2019. Using pysiaf version 0.6.1, the polynomial coefficients and other aspects of the distortion model were placed into these reference files. Previous coefficients, and the previous reference files that used them, were incorrect in that coordinates did

Data Processing Pipelines Need “Reference” Data

For a given exposure with certain combination of Detector, Filter, and Pupil settings, the same FLAT reference file is always selected for use in processing.

Alternate sets of ref files can be invoked via CRDS_CONTEXT environment variable setting.

filteroffset ---- Imager Filter Offset

flat ---- Flat Field

jwst_nircam_flat_0014.rmap

Show 200 entries Search:

DETECTOR	FILTER	PUPIL	USEAFTER	ACTIVATION DATE	REFERENCE
NRCA1	F070W	CLEAR, WLM8, WLP8	2014-06-01 00:00:00	2019-05-31 17:16:29	jwst_nircam_flat_0383.fits
NRCA1	F090W	CLEAR	2014-06-01 00:00:00	2019-05-31 17:16:29	jwst_nircam_flat_0394.fits
NRCA1	F115W	CLEAR	2014-06-01 00:00:00	2018-01-02 10:09:07	jwst_nircam_flat_0319.fits
NRCA1	F140M	CLEAR, GDHS0, GDHS60, WLM8, WLP8	2014-06-01 00:00:00	2018-01-02 10:09:07	jwst_nircam_flat_0296.fits
NRCA1	F150W	CLEAR, WLM8, WLP8	2014-06-01 00:00:00	2018-01-02 10:09:07	jwst_nircam_flat_0347.fits
NRCA1	F150W2	CLEAR, GDHS0, GDHS60	2014-06-01 00:00:00	2018-01-02 10:09:07	jwst_nircam_flat_0312.fits
NRCA1	F150W2	F162M	2014-06-01 00:00:00	2018-01-02 10:09:07	jwst_nircam_flat_0284.fits
NRCA1	F150W2	F164N	2014-06-01 00:00:00	2018-01-02 10:09:07	jwst_nircam_flat_0311.fits
NRCA1	F182M	CLEAR, WLM8, WLP8	2014-06-01 00:00:00	2018-01-02 10:09:07	jwst_nircam_flat_0348.fits
NRCA1	F187N	CLEAR, WLM8, WLP8	2014-06-01 00:00:00	2018-01-02 10:09:07	jwst_nircam_flat_0282.fits
NRCA1	F200W	CLEAR, WLM8, WLP8	2014-06-01 00:00:00	2018-01-02 10:09:07	jwst_nircam_flat_0337.fits
NRCA1	F210M	CLEAR, WLM8, WLP8	2014-06-01 00:00:00	2018-01-02 10:09:07	jwst_nircam_flat_0280.fits
NRCA1	F212N, WLP4	CLEAR, WLM8, WLP8	2014-06-01 00:00:00	2018-01-02 10:09:07	jwst_nircam_flat_0290.fits
NRCA2	F070W	CLEAR, WLM8, WLP8	2014-06-01 00:00:00	2019-05-31 17:16:29	jwst_nircam_flat_0389.fits
NRCA2	F070W	MASKRND	2014-06-01 00:00:00	2019-06-24 10:42:38	jwst_nircam_flat_0427.fits

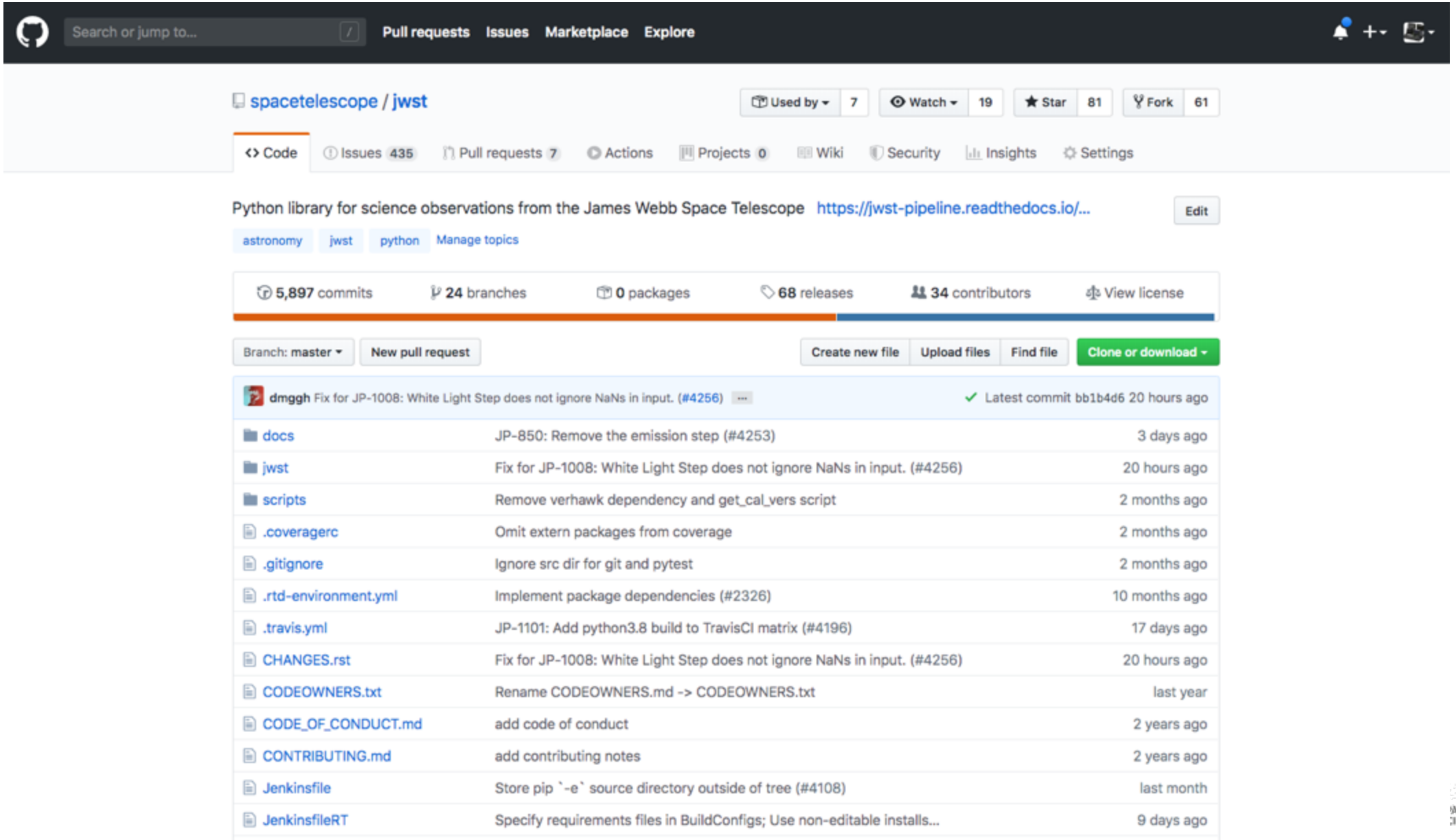
Process Logging

All relevant processing info recorded in output product meta data

```
SDP_VER = '2019_3  '           / Data processing softwa
PRD_VER = 'PRDOPSSOC-M-025'    / S&OC PRD version numbe
CAL_VER = '0.14.0  '           / Calibration software v
CAL_VCS = 'RELEASE '           / Calibration software v
```

```
        CRDS parameters

CRDS_VER= '7.4.1.2 '           / Version of CRDS file selection software used
CRDS_CTX= 'jwst_0563.pmap'     / CRDS context (.pmap) used to select ref files

R_AREA  = 'crds://jwst_nircam_area_0007.fits' / Pixel area reference file name
R_DARK  = 'crds://jwst_nircam_dark_0041.fits' / Dark reference file name
R_DISTOR= 'crds://jwst_nircam_distortion_0089.asdf' / Distortion reference file
R_FILOFF= 'N/A     '           / Filter Offset reference file name
R_FLAT  = 'crds://jwst_nircam_flat_0298.fits' / Flat reference file name
R_GAIN  = 'crds://jwst_nircam_gain_0048.fits' / Gain reference file name
R_LINEAR= 'crds://jwst_nircam_linearity_0053.fits' / Linearity reference file na
R_MASK  = 'crds://jwst_nircam_mask_0030.fits' / Mask reference file name
R_PERSAT= 'crds://jwst_nircam_persat_0020.fits' / Persistence saturation referen
R_PHOTOM= 'crds://jwst_nircam_photom_0060.fits' / Photometric reference file nam
R_READNO= 'crds://jwst_nircam_readnoise_0026.fits' / Read noise reference file n
R_SATURA= 'crds://jwst_nircam_saturation_0067.fits' / Saturation reference file
R_SPCWCS= 'N/A     '           / Spectral distortion reference file name
R_SUPERB= 'crds://jwst_nircam_superbias_0027.fits' / Superbias reference file na
R_TRPDEN= 'crds://jwst_nircam_trapdensity_0008.fits' / Trap density reference fi
R_TRPPAR= 'crds://jwst_nircam_trappars_0008.fits' / Trap parameters reference fi
R_WAVRAN= 'N/A     '           / Wavelength Range reference file name

        Calibration step information

S_WCS   = 'COMPLETE'           / Assign World Coordinate System
S_DARK  = 'COMPLETE'           / Dark Subtraction
S_DQINIT= 'COMPLETE'           / Data Quality Initialization
S_FLAT  = 'COMPLETE'           / Flat Field Correction
S_GANSCL= 'SKIPPED '           / Gain Scale Correction
S_GRPSCL= 'SKIPPED '           / Group Scale Correction
S_JUMP  = 'COMPLETE'           / Jump Detection
S_LINEAR= 'COMPLETE'           / Linearity Correction
S_PERSIS= 'COMPLETE'           / Persistence Correction
S_PHOTOM= 'COMPLETE'           / Photometric Calibration
S_RAMP  = 'COMPLETE'           / Ramp Fitting
S_REFPIX= 'COMPLETE'           / Reference Pixel Correction
S_SATURA= 'COMPLETE'           / Saturation Checking
S_SUPERB= 'COMPLETE'           / Superbias Subtraction
```

JWST Calibration Pipeline Package is Open Source



And straight-forward to install

uires Python 3.5 or above and a C compiler for dependencies.

tion

package can be installed into a virtualenv or conda environment via `pip`. We recommend creating a fre
nt with only python installed. Via conda:

```
reate -n jwst_env python=3.7.4
ctivate jwst_env
```

g for end-users

released (tagged) version, you can install directly from Github. To install tagged release `jwst 0.14.0`:

```
tall git+https://github.com/spacetelescope/jwst@0.14.0
```

development version (from `master`) can also be installed from Github:

```
tall git+https://github.com/spacetelescope/jwst
```

articular commit hash:

```
tall git+https://github.com/spacetelescope/jwst@3f03323c
```

Dependency Management & Testing

- Third-party package changes can affect results
- Use conda environment files to lock down (pin) dependencies
  - conda create --n [env_name] --file https://ssb.stsci.edu/releases/jwstdp/0.14.1/conda_env_dump_osx-stable-deps.txt

- Unit tests (embedded in code) and regression tests help maintain consistent results
  - Test against
    - Pinned versions of 3rd-party libraries
    - Released (stable) versions of 3rd-party libraries
    - Latest development (github master branch) versions of 3rd-party libraries


EXPANDING THE FRONTIERS OF SPACE ASTRONOMY