The JWST pipeline will never automate everything for everybody, but it should be capable of doing most things for most people.

STScI | SPACE TELESCOPE SCIENCE INSTITUTE

EXPANDING THE FRONTIERS OF SPACE ASTRONOMY

JWST Pipeline Overview

David R. Law

JWST Pipeline Lead

January 11, 2025

AAS Pipeline Workshop

Orion Nebula: ESA/NASA/CSA

Workshop schedule

9:00-9:05: Welcome and Introduction

9:05-9:35: JWST pipeline overview presentation (David)

9:35-10:30: MIRI/NIRCam/NIRISS imaging notebooks

10:30-11:00 Break

11:00-12:00: MIRI and NIRCam known issues and ongoing calibration efforts.

12:00-13:00 Lunch

13:00-14:00 NIRISS and NIRSpec known issues and ongoing calibration efforts.

14:00-15:00: NIRSpec slit spectroscopy notebook

15:00-15:15 Break

15:15-16:15: MIRI MRS notebook

16:15-17:00: General discussion, Q&A, closing survey.

Who should I talk to today?

David Law
Pipeline lead, MIRI team

Chris Hayes
Deputy lead, NIRSpec team

Nadia Dencheva
Pipeline developer

Brett Graham
Pipeline developer

Steph LaMassa
NIRISS team

Alicia Canipe
NIRCam team

Macarena Garcia Marin
Project Scientist

Mic Bagley
Project Scientist

Stacey Bright
Project Scientist

Pipeline basics: software + reference files

Pipeline code

- Shared across all instruments: https://github.com/spacetelescope/jwst
- Documentation: https://jwst-docs.stsci.edu
- Dependencies on other STScI packages, astropy, numpy, etc.
- Mostly python, some C bindings for speed
- Highly modular; can skip, modify, or add steps
- Supported for Mac OS, Linux

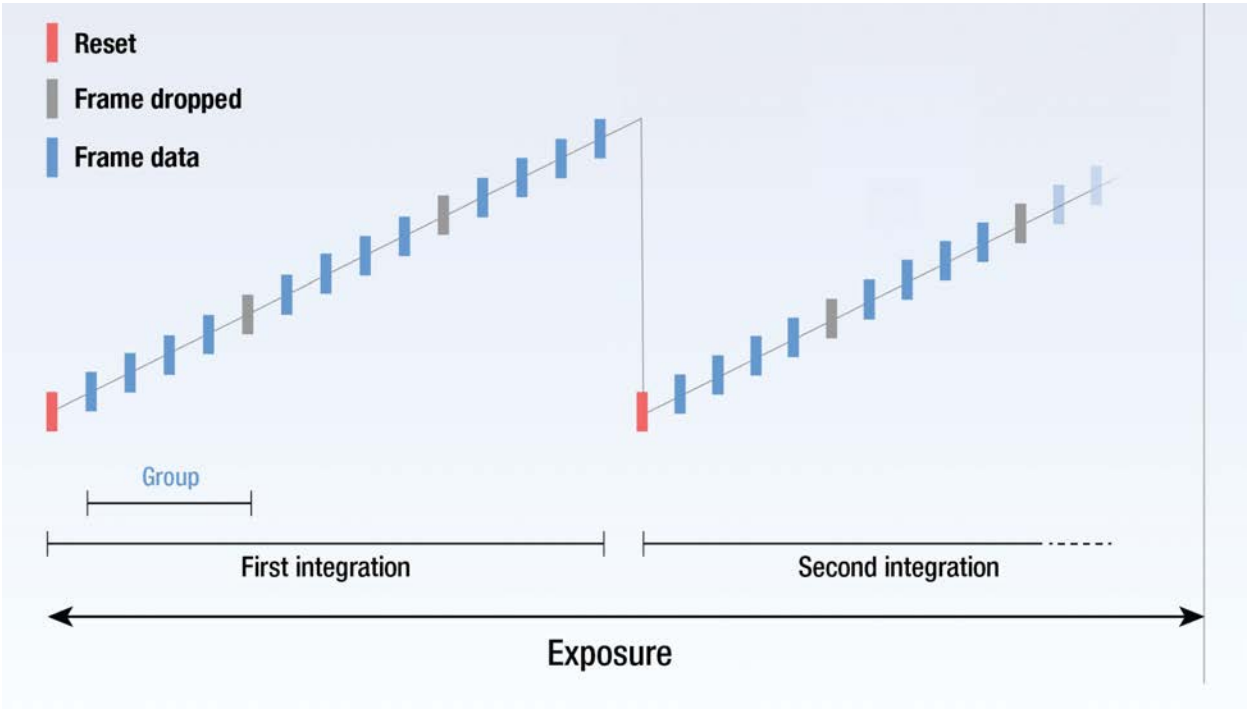
Reference files

- Many pipeline steps rely on reference files; will be retrieved automatically as needed
- These are stored in the CRDS server (https://jwst-crds.stsci.edu/)
- Each pipeline version corresponds to a fixed set of reference files

Pipeline notebooks put this together

- https://github.com/spacetelescope/jwst-pipeline-notebooks

Pipeline basics: uncalibrated data



JWST uses up-the-ramp sampling to non-destructively read each detector array

Raw data products are thus 4-D; ncolumns x nrows x ngroups x nintegrations

See https://jwst-docs.stsci.edu/understanding-exposure-times

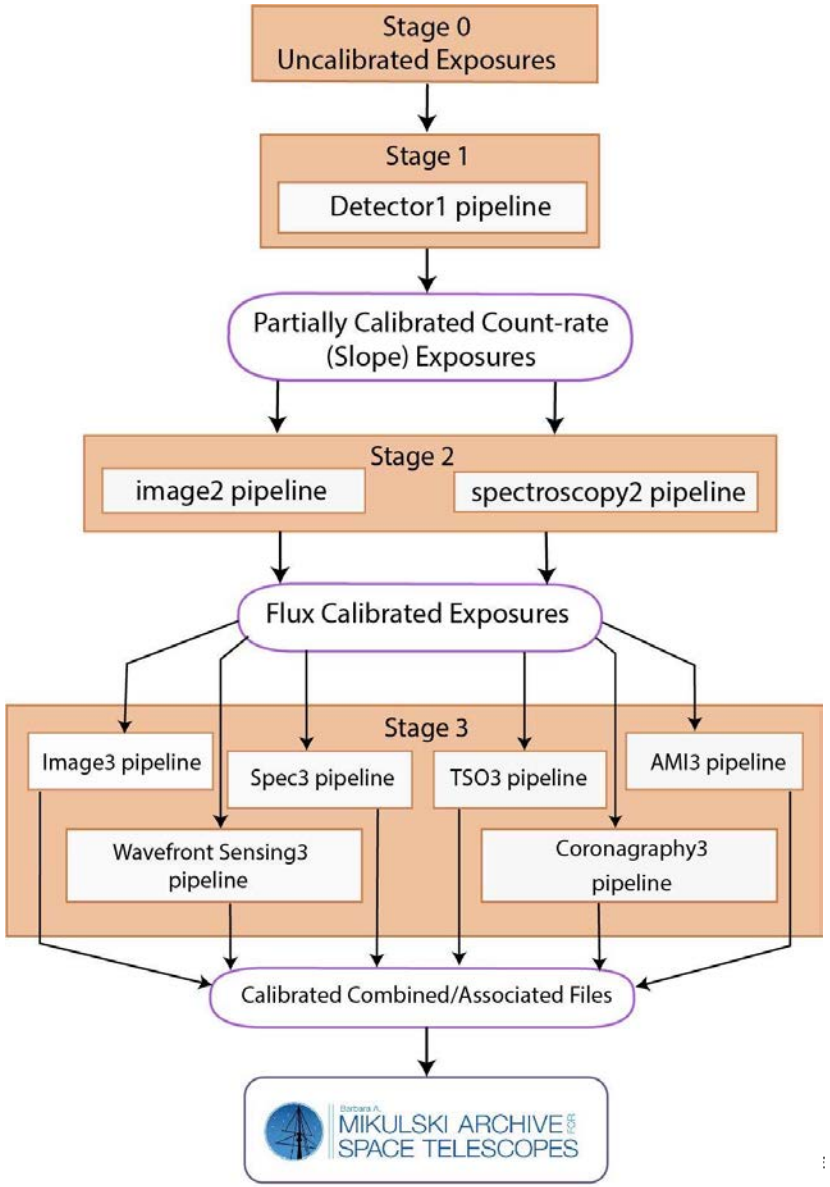
Pipeline basics: processing stages

Input: Uncalibrated ramp data (uncal.fits)

Stage 1: calwebb_detector1
Remove detector effects, fit ramps to 4D slopes.
Output is 2D detector images in units of DN/s (rate.fits)

Stage 2: calwebb_spec2, calwebb_image2
Calibration; split by imaging vs spectroscopy.
Output is calibrated 2D detector images in units of MJy/sr (cal.fits)

Stage 3: calwebb_image3, calwebb_spec3, calwebb_coron3, etc
Split by mode; combine multiple exposures
Output is calibrated 2D images (i2d.fits), 3D IFU cubes (s3d.fits), 1D spectra (x1d.fits), etc.



Development follows a quarterly schedule

A ‘build’ is a released and tagged set of software including the core ‘jwst’ package and related dependencies.

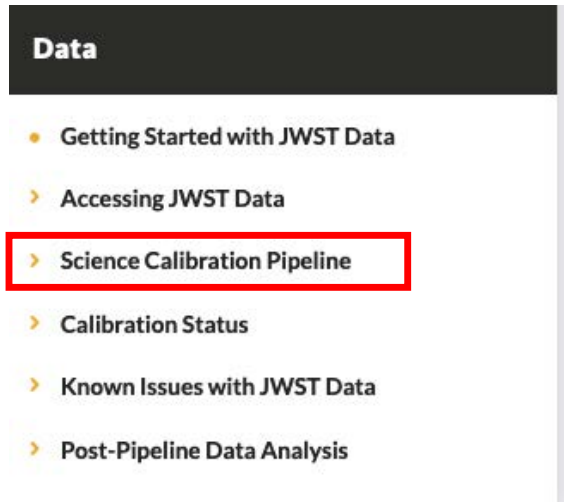
See JDox for the latest table

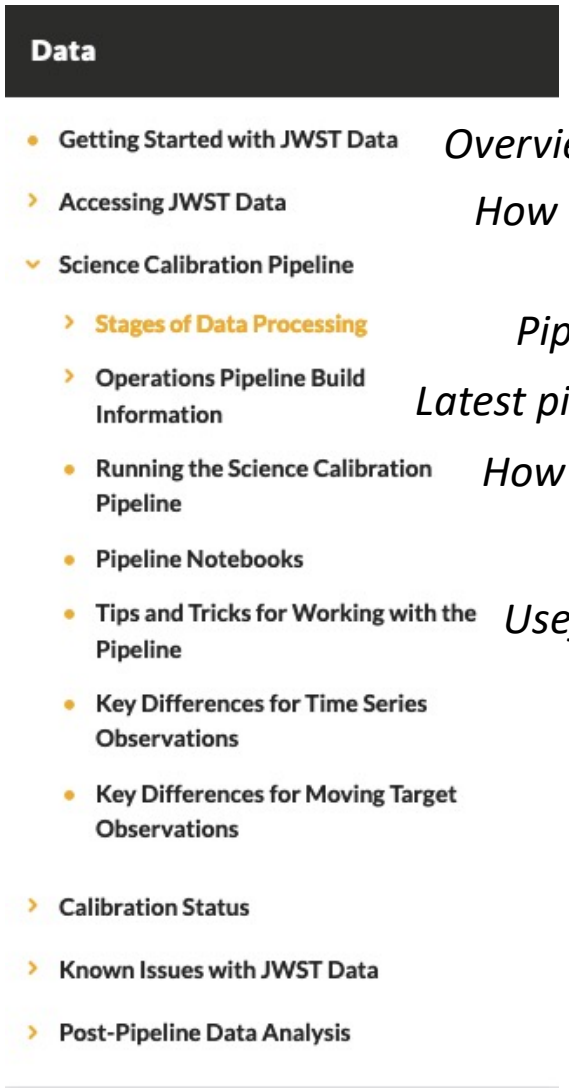
Operations Pipeline Build	Science Calibration Pipeline Version (jwst)	Build Status	SDP_VER	Candidate Released	Operations Installation	CRDS context
11.3	TBD	Development	TBD	Estimated Spring 2025	Estimated Summer 2025	TBD
11.2	1.17.1	Candidate	TBD	2025-01-03	Estimated Mar 2025	1321
11.1.1	1.16.1	Operations	2024.3.1	2024-09-20	2024-12-03	1303
11.0.1	1.15.1	Archived	2024.2.1	2024-06-26	2024-08-27	1293

Build lifecyle:

- Development
  - Changing daily, not recommended for non-developers. ‘jwst’ main branch.
- Candidate
  - Development finished, working on testing and documentation. Available for interested users.
- Operational
  - Default recommendation for users. Used to produce MAST data products.
- Archived
  - Old versions that are no longer current or recommended

How to get the latest information





See https://jwst-docs.stsci.edu

Overview of the documentation and where to find information

How to download JWST data

Pipeline steps and algorithms

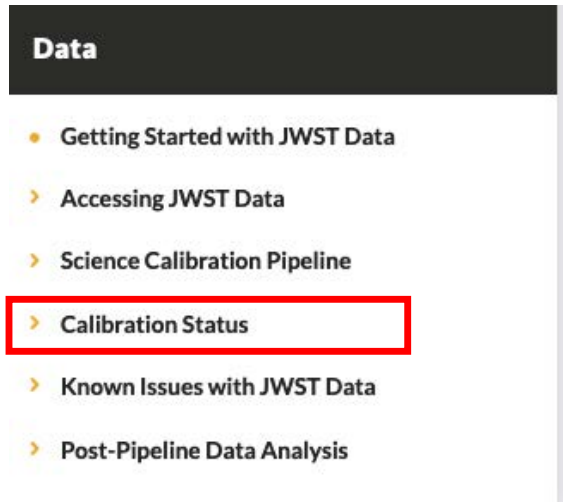
Latest pipeline updates

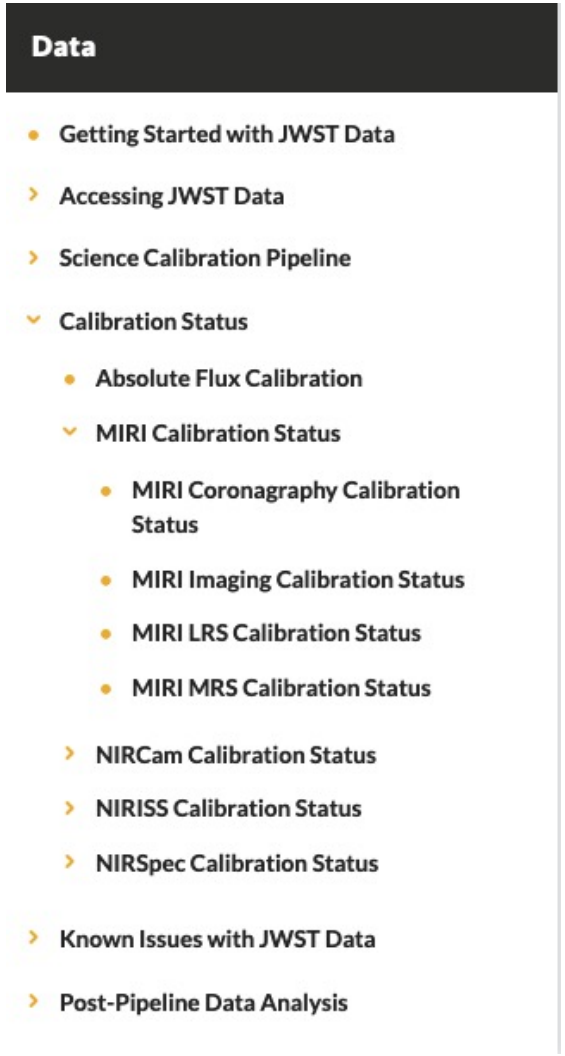
How to run the pipeline

Useful stuff I keep needing to look up!

8

How to get the latest information



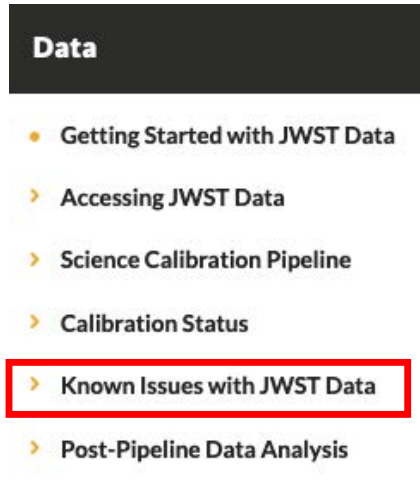


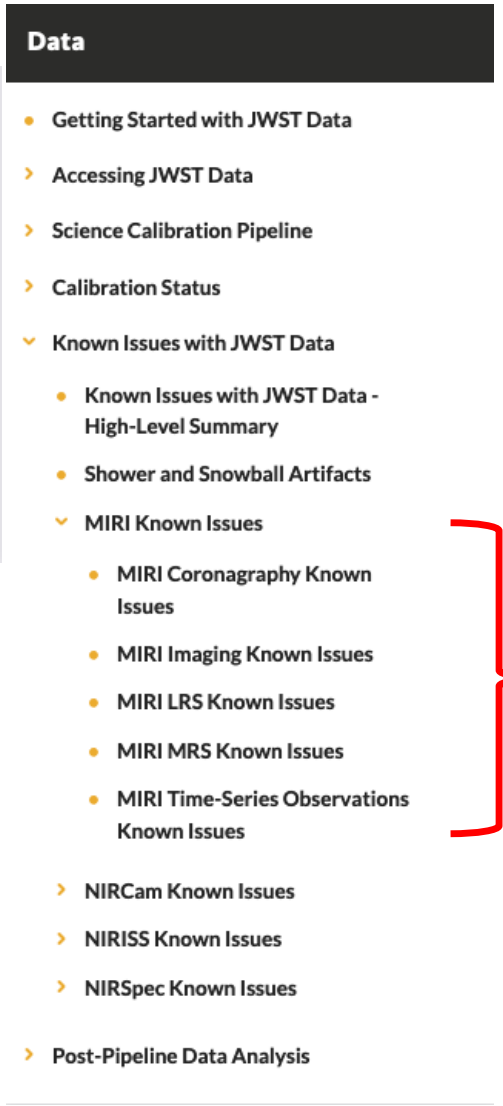
See https://jwst-docs.stsci.edu

MIRI Information

9

How to get the latest information





MIRI Information

See https://jwst-docs.stsci.edu

10

What has changed recently? (A very abbreviated list)

Build 11.1 (jwst version 1.16.1)

- Added a 1/f noise correction step to calwebb_detector1
- Improvements to runtime and memory management
- Improved NIRISS SOSS spectral extraction using PASTASOSS package
- Fixed some NIRSpec flux calibration issues

Build 11.2 (jwst version 1.17.1)

- Improvements to NIRSpec outlier detection
- Automated flagging of cosmic ray showers for MIRI MRS
- Improvement to MOS and WFSS background subtraction
- Fixed a bug in mosaic image alignment from build 11.1

Why run the pipeline yourself?

MAST data products use default parameters to give reasonable results for ALL observing programs (Jupiter to cosmology, time-domain exoplanets to nearby galaxy mosaics, etc)

Many data products can be improved for specific science cases by using non-default settings!

Jupyter notebooks giving examples of how to run the pipeline

https://github.com/spacetelescope/jwst-pipeline-notebooks

Currently examples for MIRI/NIRCam/NIRISS imaging, MIRI MRS spectroscopy

We'll be using modified versions of these notebooks today

- Tweaked to use offline cache of reference files
- Fixed to build 11.0 (jwst 1.15.1), while github will continually be updated
- Additional draft notebook for NIRSpec slit spectroscopy

Installing the pipeline

1) Install anaconda python. This will allow you to set up a unique python environment for the jwst pipeline.

2) Using a terminal window, create a new conda environment and install the pipeline:
conda create -n jwst1.15.1 python=3.11
conda activate jwst1.15.1
pip install jwst==1.15.1
pip install gwcs==0.21.0 (new requirement to deal with breaking change in gwcs)

3) Install other necessary python packages
conda activate jwst1.15.1 (if it isn't already active)
pip install jupyter nbclassic astroquery jdaviz

4) (Optional for this workshop, but required in general)
Set environment variables (in your bash profile or equivalent) linking to the CRDS reference file server.

export CRDS_PATH=<locally-accessible-path>/crds_cache/jwst
export CRDS_SERVER_URL=https://jwst-crds.stsci.edu

Installing the workshop data packs

JWST data takes up a lot of space!

- Some reference files are 6 GB each
- Raw and processed data from multiple detectors can often exceed 10 GB
- Free up some disk space before working with JWST data!

These packs contain pipeline notebooks, example JWST data from real observing programs, and the necessary calibration files.

Now let's open the data packs!

MIRI Imaging:
https://stsci.box.com/s/mxtc869djgva0qbkn6ixwdjrg84uf1t7
tarball: 2.7 GB
total: 7.1 GB
runtime: 32 seconds

MIRI MRS:
https://stsci.box.com/s/ojvhywmvukcgq5la6wcsl1pet6ghg9bp
tarball: 8.5 GB
total: 23.3 GB
runtime: 1433 seconds

NIRCam Imaging:
https://stsci.box.com/s/441q2ogpkh8n8562gbn8em0op8mncz07
tarball: 10.0 GB
total: 22.1 GB
runtime: 232 seconds

NIRISS Imaging:
https://stsci.box.com/s/0dk7a8vvhs6k2gwh2eir2dc6birf4ny1
tarball: 3.2 GB
total: 10.8 GB
runtime: 922 seconds

NIRSpec Slit:
https://stsci.box.com/s/rxnt003tuqlkv71jyedqvnydjzhowpec
tarball: 1.0 GB
total: 1.4 GB
runtime: 54 seconds