STScI | SPACE TELESCOPE SCIENCE INSTITUTE

EXPANDING THE FRONTIERS OF SPACE ASTRONOMY

Adding Science Validation to the JWST Calibration Pipeline.

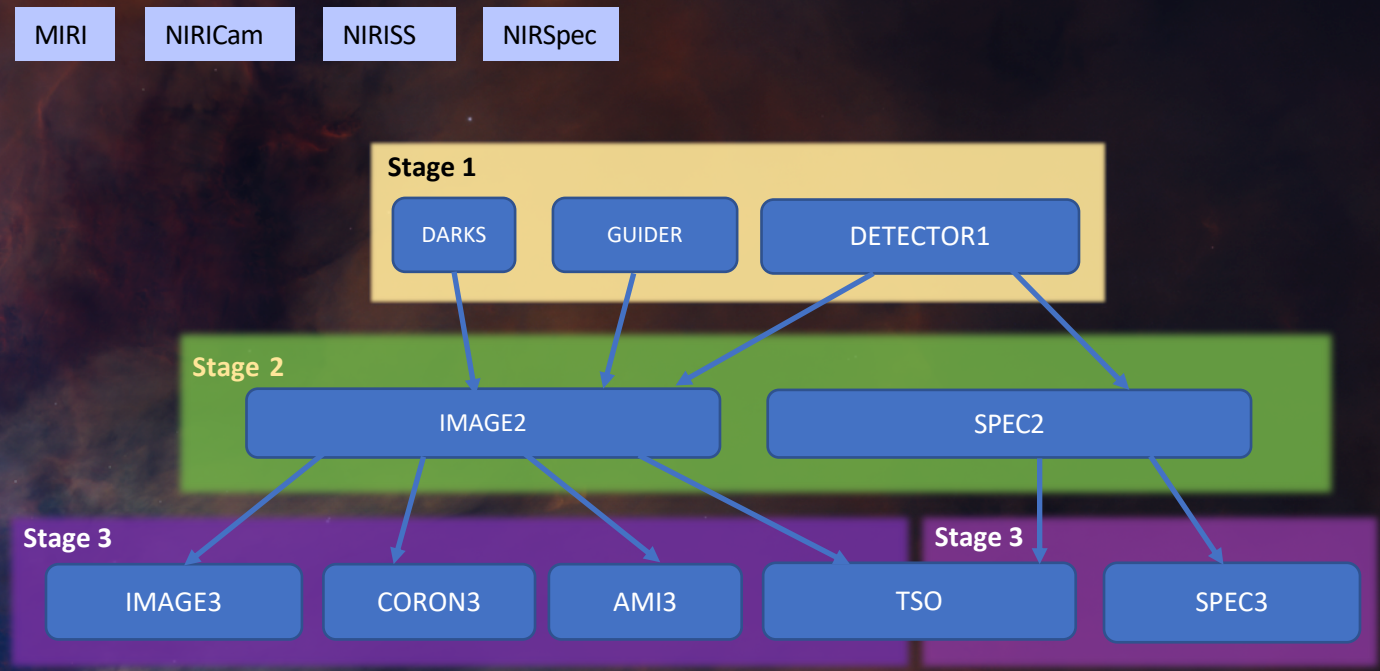
Rosa I Diaz & Macarena Garcia Marin

ADASS XXVIII

November 12, 2018

The JWST Calibration Pipeline

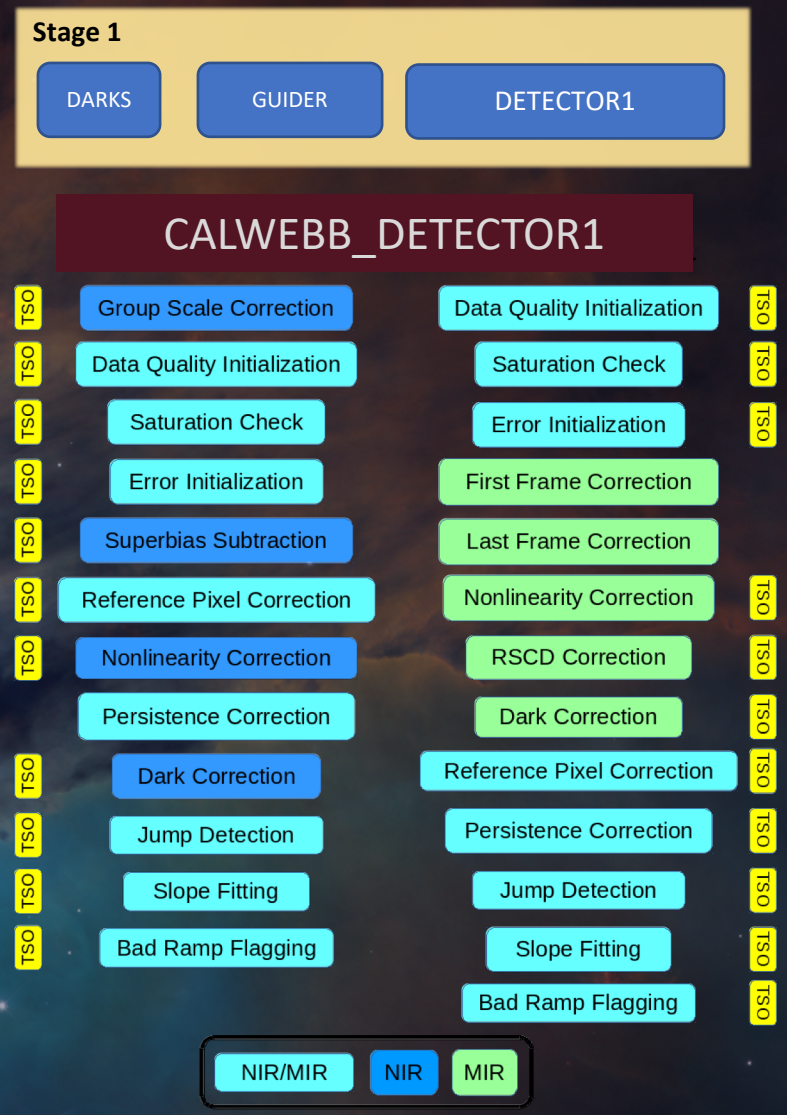
- Is a Python software suite that automatically process the data taken by all the JWST instruments and observing modes
- It produces both fully calibrated individual exposures and high level data products (dithers, mosaics, different detector, etc.).
- It is divided in tree different stages. Each of which can be subdivided according to the type of observation
- Stage 1 that will include all the steps that are common to all type of data
- stage 2 calibrates the individual slope images or spectra
- Stage 3 combines the slope images or spectra into a single product.
- Each stage is further subdivided in steps that are executed according to the type of instrument or observation



The JWST Calibration Pipeline

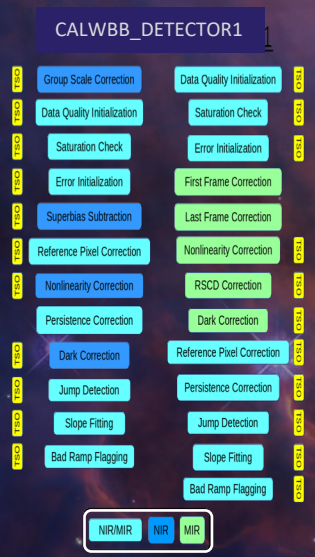
- Is a Python software suite that automatically process the data taken by all the JWST instruments and observing modes
- It produces both fully calibrated individual exposures and high level data products (mosaics, extracted spectra, etc.).
- It is divided in tree different stages. Each of which can be subdivided according to the type of observation
- Each stage is further subdivided in steps that are executed according to the type of instrument or observation

MIRI | NIRICam | NIRISS | NIRSpec

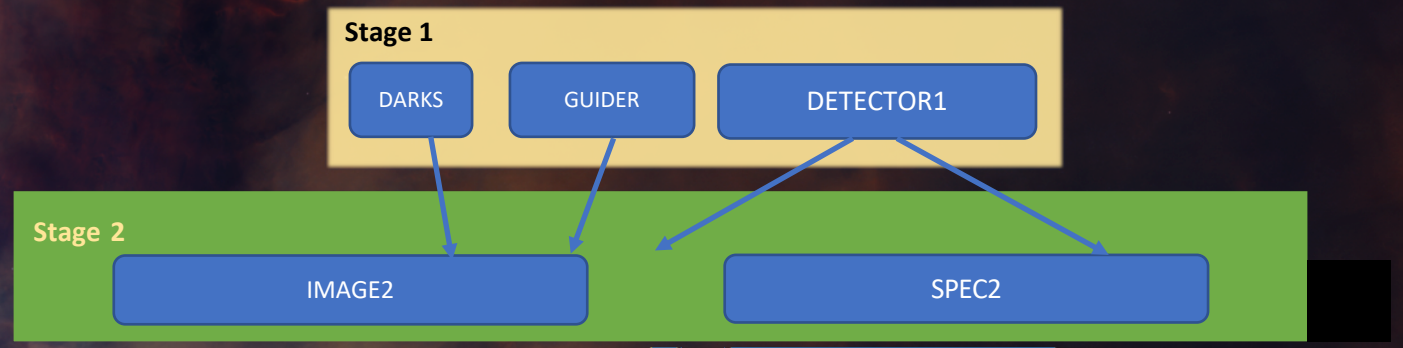


The JWST Calibration Pipeline

- Is a Python software suite that automatically process the data taken by all the JWST instruments and observing modes
- It produces both fully calibrated individual exposures and high level data products (mosaics, extracted spectra, etc.).
- It is divided in tree different stages. Each of which can be subdivided according to the type of observation
- Each stage is subdivided in steps that are executed according to the type of instrument or observation



MIRI | NIRICam | NIRISS | NIRSpec



TSO WCS Information

Background Subtraction

TSO Flat Field Correction

TSO Flux Calibration

TSO Rectified 2D Product

TSO WCS Information

Background Subtraction

Imprint Subtraction — NIRSpec MSA & IFU

MSA Failed Open Flagging — NIRSpec MSA & IFU

Subwindow Extraction — NIRSpec MSA/Fixed Slit, NIRCam/NIRISS WFSS, NIRIS SOSS

TSO Flat Field Correction

Point vs Extended Decision

Stray Light Subtraction — MIRI MRS

Fringing Removal — MIRI MRS

Path-Loss Correction

TSO Flux Calibration

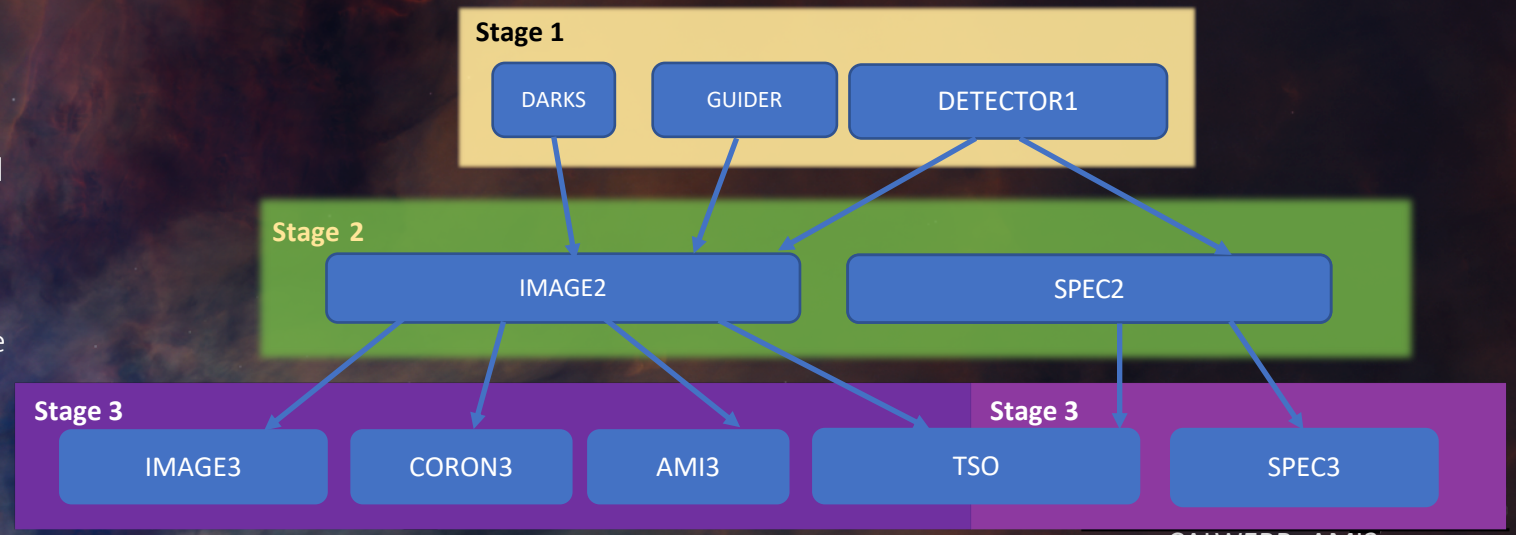
TSO Rectified 2D/3D Product

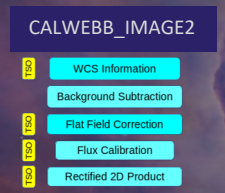
NIR/MIR | NIR | MIR

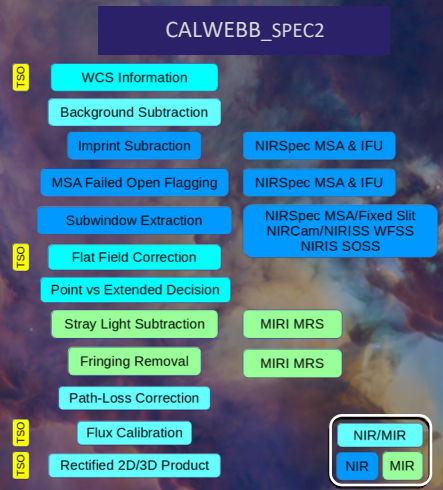
The JWST Calibration Pipeline

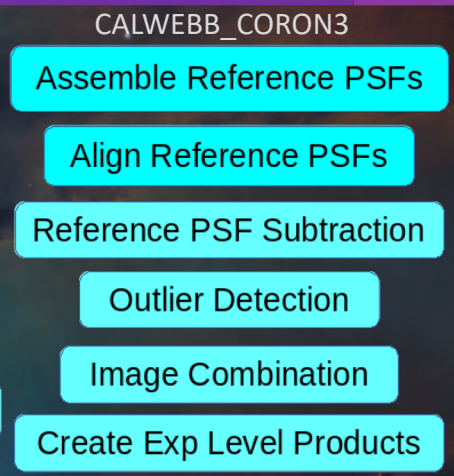
- Is a Python software suite that automatically process the data taken by all the JWST instruments and observing modes
- It produces both fully calibrated individual exposures and high level data products (mosaics, extracted spectra, etc.).
- It is divided in tree different stages. Each of which can be subdivided according to the type of observation
- Each stage is subdivided in steps that are executed according to the type of instrument or observation

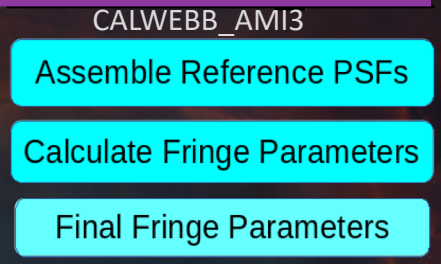
MIRI | NIRICam | NIRISS | NIRSpec











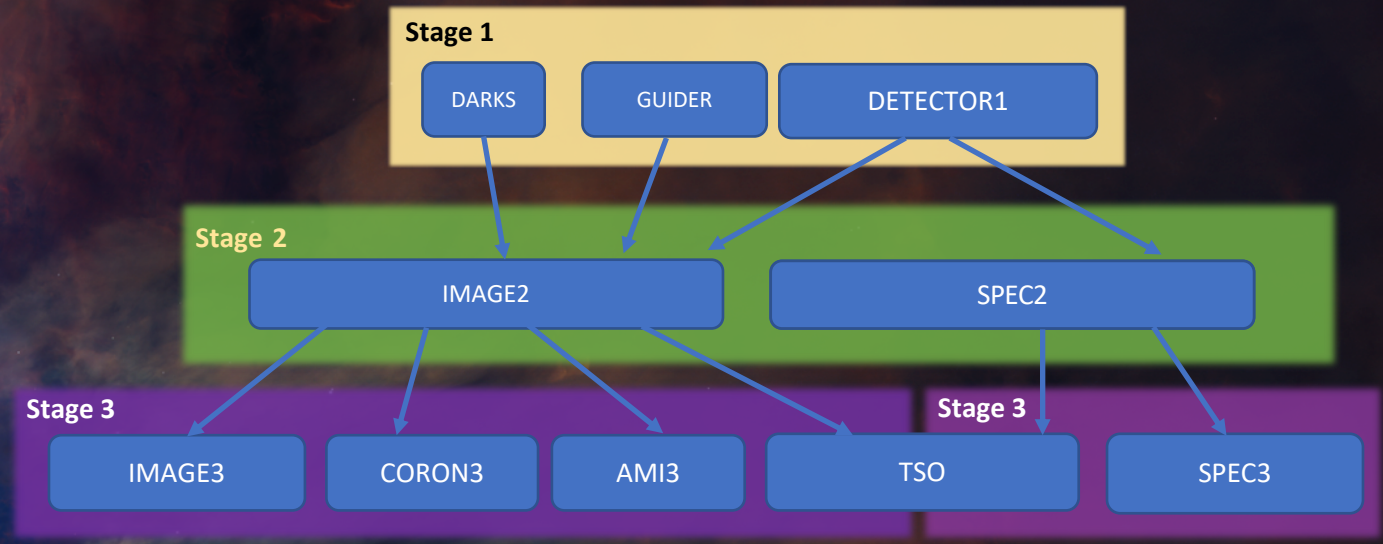
The JWST Calibration Pipeline

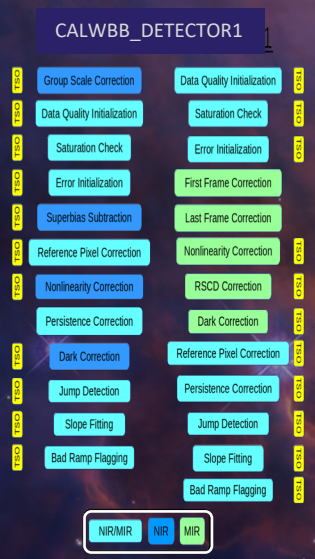
- Is a Python software suite that automatically process the data taken by all the JWST instruments and observing modes
- It produces both fully calibrated individual exposures and high level data products (mosaics, extracted spectra, etc.).
- It is divided in tree different stages. Each of which can be subdivided according to the type of observation
- Each stage is subdivided in steps that are executed according to the type of instrument or observation

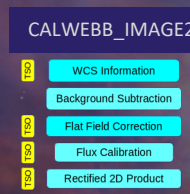
The JWST Calibration Pipeline

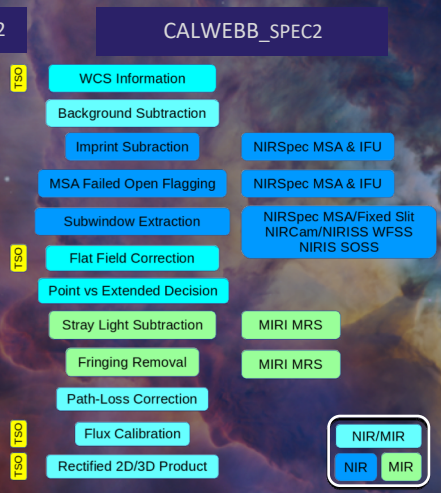
- Is a Python software suite that automatically process the data taken by all the JWST instruments and observing modes
- It produces both fully calibrated individual exposures and high level data products (mosaics, extracted spectra, etc.).
- It is divided in tree different stages. Each of which can be subdivided according to the type of observation
- Each stage is subdivided in steps that are executed according to the type of instrument or observation

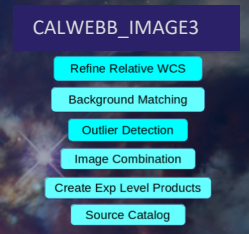
MIRI | NIRICam | NIRISS | NIRSpec

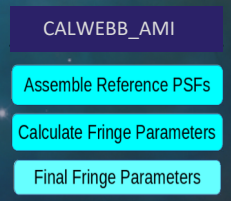


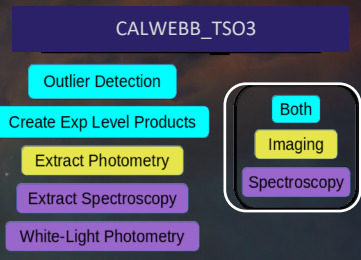


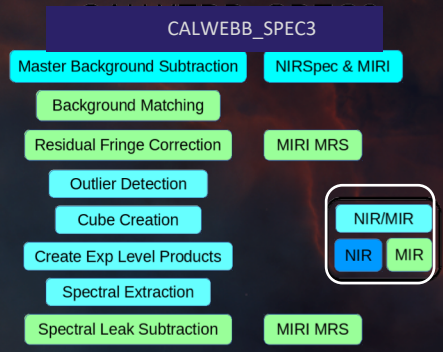












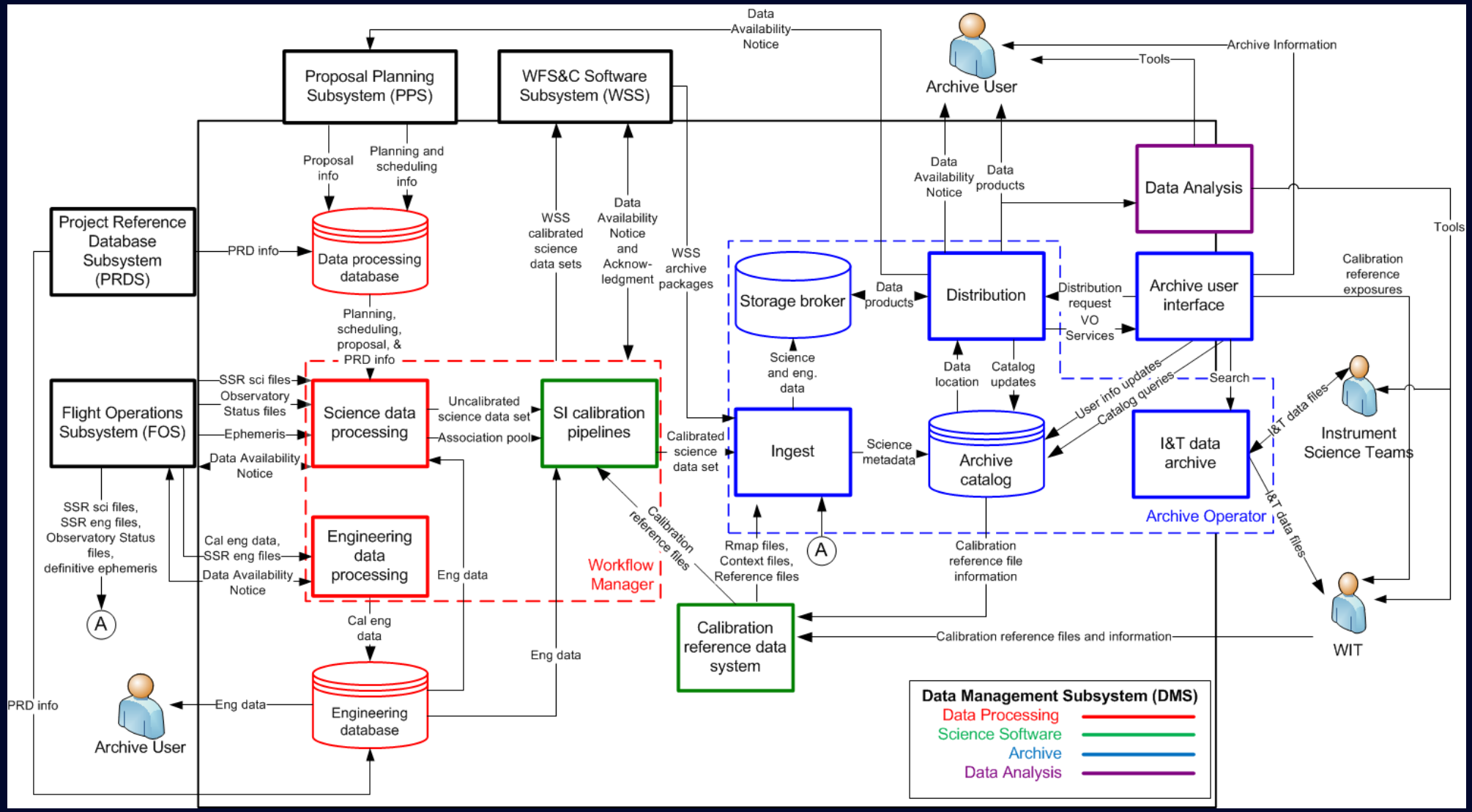
Why is it important to involve the science staff?

Careful scientific validation and verification are necessary to:

- Assess that the Cal Pipeline has full coverage for the calibration of all the data that JWST will take
- Compare different optional parameters of the software and validate differences
- Determine compatibility and content of reference calibration data with software
- Determine whether the calibration software correctly implements the specified algorithms and options.
- Determine how well the Calibration Pipeline works for all the types of science that JWST will take
  - Edge cases
  - With selected default values and thresholds
  - Reference data used
- Assess whether the defined algorithms produce products up to the standards defined by the error budget

What can Instrument Scientists do?

- Initially the process envisioned for the I&T team to fully validate the JWST pipeline. This quickly became a challenge:

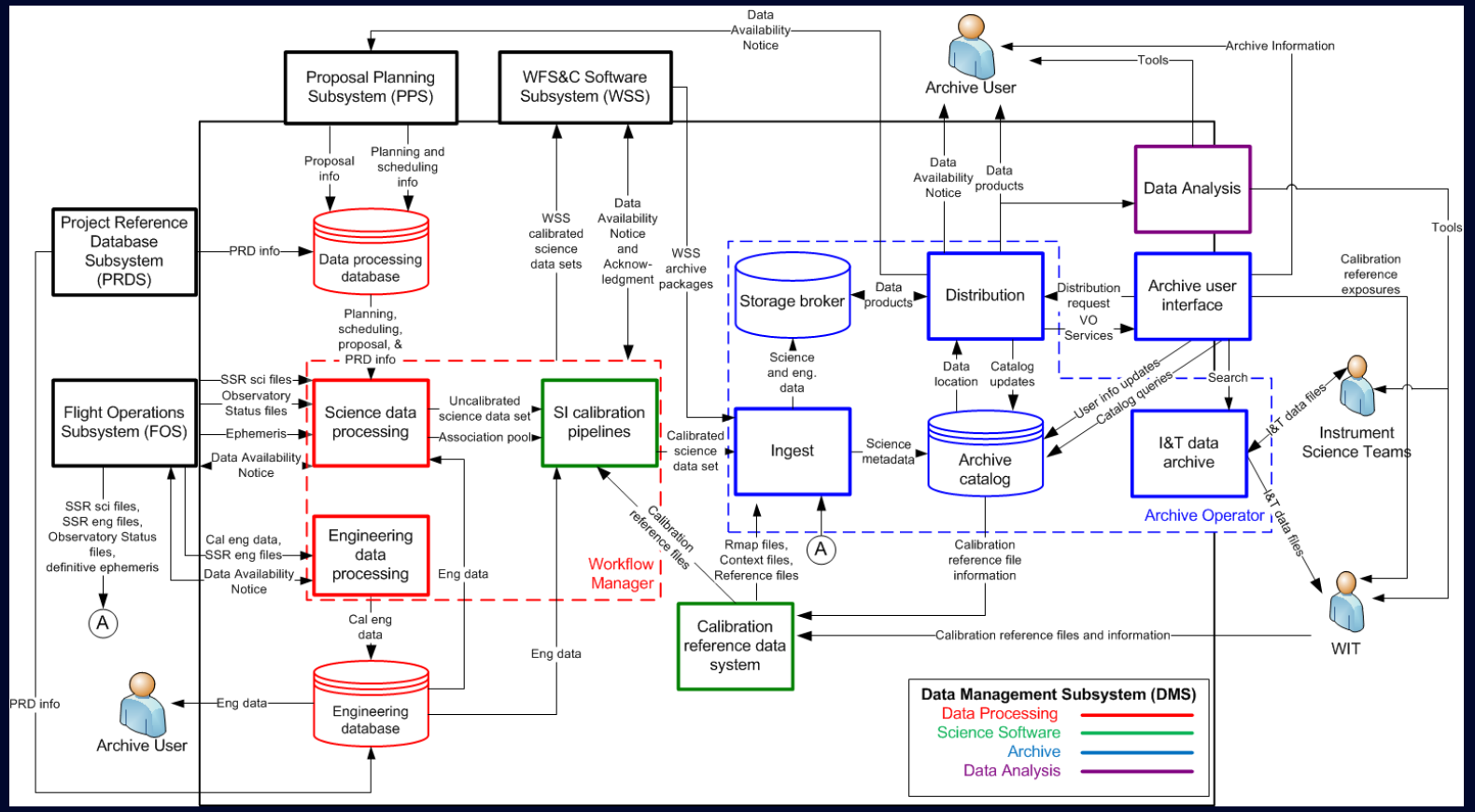


What can Instrument Scientists do?

- Initially the process envisioned for the I&T team to fully validate the JWST pipeline. This quickly became a challenge:
- Conflicts with schedule (engineers and science team) and different scope and goals:
  - Software Engineering and development requires verification of the full set of instruments and observation modes. It is also concerned about the design, performance, optimization, and maintainability of the system.
  - Scientific validation cares about validating the accuracy of the data for a large dynamic range of science cases, simplicity of use, exploration, stability, and good documentation.

We decided to adopt a more collaborative approach:

- I&T performs full JWST Pipeline verification
- Instrument Science Teams perform science validation.



Science Validation Plan

- Incremental science validation:
  - Science validation can be done as steps become available
  - some steps can be validated by a single instrument team and as resources become available
  - Validation can be done with only a subset of possible cases
- First check that calibration pipeline meets science requirements defined at the start of the mission planning (Part 1)
  - Validate the Calibration Pipeline does what the scientists defined via algorithms and functionality
  - Can be done using simulated or cryo test data and ground or dummy reference data
- Determine the accuracy and quality that the Pipeline can achieve (Part 2).
  - To what extent the selected algorithms meet the error budget and how these vary for different types of data and science cases
  - Can be done with simulated or cryo test data but also needs inflight data

Science validation plan goals and status

- We want to be able to do many of these test with each build
- We want uniformity on testing across instrument teams
- We want to be able to do it fast and minimize resources
- Within our plan, we are defining a series of validation tests that can be scripted in a tool to help us speed up the process of validation
- Currently we have about 426 tests for the calibration pipeline (these will increase as we revise the plan):
    - 192 are considered basic tests
    - 127 classified as part 1
    - 107 classified as part 2
    - 212 of all test can be done via a computational algorithm (probably more will be added later)
    - 101 already have some sort of script developed by one or several of the instrument teams.
- We have developed simulated data or identified cryo or ground test data to use for the validation

Science validation plan goals and status (cont)

- From those that can be coded, we are re-evaluating the full set to classify them in the following categories:
  - Test suitable to become unit test (201)
    - Currently working with the development team to incorporate these test within the Calibration Pipeline code.
  - Test suitable for regression testing (11)
    - Incorporate these test within an automated framework and web interface for quick analysis and evaluation of the results by the instrument teams
- We are currently holding walkthroughs for each of the most complex steps in order to have a better insight on the code and be able to design the science validation tests.

Status, future plans, and goals

- Have the baseline validation plan by the spring of 2019
- Have a full validation plan by the end of 2019
- Goal is to migrate any test that does not require human interaction/interpretation to the Calibration pipeline software
- Those that need human interpretation and should be repeated with each build will be part of a Calibration Pipeline Testing Tool
- We are building a complete regression test suite
  - tests that cover the full range of science cases planned with JWST
  - Share regression test data for use by developers of the Calibration Pipeline software, scientist, and any other teams developing tools to support analysis of JWST data.

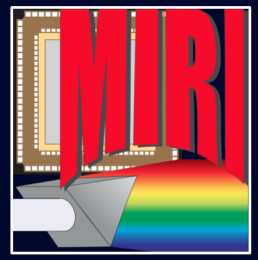
Our approach to validation

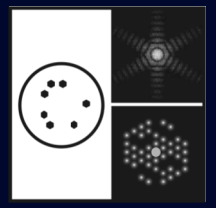
- Scientist and developers working together to test the calibration software:

  "The combined effects is greater than the sum of their separate effects."

  - Sharing code and ideas among instrument teams to support validation
  - Sharing code and test ideas with developers
  - Compare testing results among the different teams
  - Flexibility in schedule
  - Makes it easier to understand the needs of developers

- Challenges
  - Requires openness among these two groups to feedback and acknowledgement of experience.
  - Requires for scientist to learn about development practices followed by the SCSB team
  - Requires for SCSB to fell comfortable allowing scientists to actively participate in the software development process.
  - Requires to get the best of a team, in spite of different points of view and priorities.





AMI

STScI WGs and Teams

- WFSC Team
- TSO WG
- Coronagraphic WG
- Calibration Pipeline WG (algorithm definition)

SCSB
Science Instrument Calibration Software Branch

DMS
Data Management System