JSTUC Data Analysis Advisory Group

Final Report

December 11, 2020

George Becker (chair) (UC Riverside)

Jennifer Andrews (U of Arizona)

Zach Berta-Thompson (U Colorado Boulder)

Marusa Bradac (UC Davis)

Stephane Charlot (IAP/CNRS-Sorbonne U)

Casey Papovich (Texas A&M University)

Johan Richard (CRAL)

Mike Wong (UC Berkeley)

Previous Recommendations (June 2020)

1. Enhance communication between STScI and community teams
   - Amaya Moro-Martin facilitating communication with ERS teams
   - Meeting between DAT team and ERS teams scheduled for December
2. Increase access to simulated data
   - STScI working to organize existing simulated data, run training sessions
   - NIRSpec IDT/GTO team processing simulations on a per-request basis, running training sessions
3. Prioritize DA tools that address JWST-specific needs
4. Facilitate existing DA tools that are either widely used or difficult to replace

Data Analysis Tools Survey

Since June 2020, the DAAG has moved its focus to the broader landscape of DA tools and resources that will be needed by community groups.

Guiding principle is **rapid and flexible** exploitation of JWST data.

DA Tools Survey sent out to ERS, GTO teams, and community groups

- Goals
  - Develop a more holistic view of how teams expect to approach their JWST data
  - Assess the utility of published STScI DA tools (Jupyter notebooks, visualization tools, etc)
  - Identify specific new tools or resources that may be needed
  - Gauge expectations for existing tools
- Complements STScI's 2019 JWST Data Analysis User Survey and ongoing engagement with ERS teams
- Good response rate
  - 17 replies, mainly from ERS teams
  - Broad scope in terms of science (solar system to extragalactic) and observing modes

Survey Results: Most relevant STScI Tools

Among the tools listed on the JDox website above, which, if any, seem most relevant to your program?

- Includes only tools published on the JDox site as of Nov 2020
- Strong enthusiasm for:
  - Visualization tools: jdaviz (specviz, cubeviz, mosviz)
  - Spectral extraction tools: MOS Optimal Extraction (NIRSpec), IFU Cube Analysis (NIRSpec, MIRI), BOTS (NIRSpec); grizli also mentioned w/r/t NIRISS
  - Photometric tools (NIRCam)
  - Astropy utilities: imexam, photutils, specutils, etc.
- Use of DRP frequently mentioned

Takeaway: Highest value placed on tools for visualization and producing science-ready data products that can be fed to other analysis tools.

Survey Results: Utility of STScI Tools

How well do the tools listed on the JDox website above appear to meet the (post-pipeline) data analysis requirements for your JWST program?

- Wide range of responses, as expected for new tools not yet widely deployed
- Many respondents emphasized that the STScI tools would only be part of their larger toolkit.

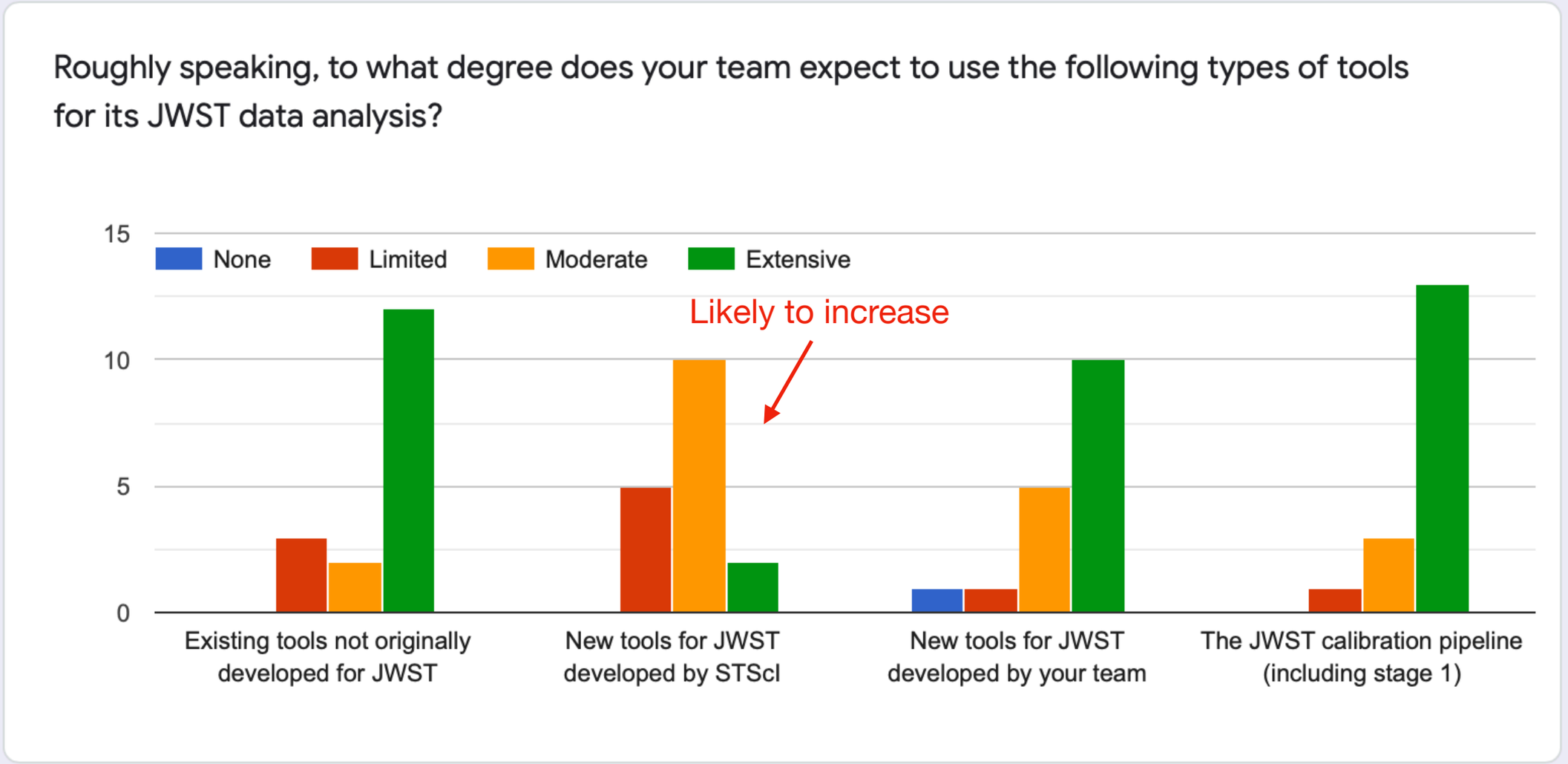
Survey Results: Needed Tools

Are there critical tools missing from this list that your team would like to see developed?

- Diverse list
  - Tools for non-redundant masking of NIRISS
  - Complicated background subtraction tools
  - Mosaic tools for Solar System targets (MIRI and NIRSpec)
  - ASDF reader for C/C++
  - NIRCam transiting exoplanet notebook
  - Generalized spectra extraction tools for remaining instruments
  - Drizzle capability for IFU data cubes
  - Tools to inter-compare NIRSpec IFU and MIRI MRS datacubes
  - Multi-band photometry with PSF corrections
- List will be communicated to STScI DAT team, may inform upcoming discussions with ERS teams.

Takeaway: Most requests are for tools that address JWST-specific data characteristics.

Survey Results: Mix of Tools



Takeaway: Even with a strong uptake of new tools from STScI, teams expect to rely heavily on existing tools or new tools they will develop themselves for JWST. Interacting with the DRP is also likely to be critical for many teams.

Survey Results: Critical Existing Tools

> What existing (legacy) tools does your team expect to be critical for analyzing its JWST data? These may include tools for visualization, photometric or spectroscopic measurements, spectral modeling, time series analysis, light curve modeling, etc.

- Very long list
- Some tools apply to “fully processed” data.
  - SED fitting, spectral modeling, photo-z’s, time series analysis, etc.
- Others will need to be adapted to JWST data.
  - Visualization: ds9, QFitsView
  - Slitless spectral extraction: grizli (WFC3 —> NIRISS)
  - Source extraction and photometry: SExtractor, DAOPHOT, DOLPHOT, T-PHOT
  - Image reconstruction for non-redundant masking
  - Tools to provide backplane data to spacecraft observations (through SPICE) and ground-based observations

Takeaway: Teams expect to rely heavily on a wide range of tools that will need to be adapted to JWST.

Survey Results: Access to Simulated Data

If there are existing tools critical to your team that will require specific JWST compatibility, is it a high priority for your team to test these tools with simulated data? If so, is your team able to generate or obtain the simulated data that are needed?

- Most teams report a need for simulated data for software testing
- Several teams report success with Mirage (NIRCam and NIRISS) and MIRIsim
- A few have obtained NIRSpec simulated data through the NIRSpec IDT/GTO team.

Takeaway: Reinforces the need for a well-organized **repository of simulated data** maintained by STScI, and training on simulation tools. This is currently underway.

NIRSpec simulations: Through an effort coordinated by STScI and with support from the NIRSpec STScI group, the NIRSpec IDT/GTO team is training the interested ERS teams to prepare simulation inputs and is working to generate the corresponding simulated data.

Survey Results: Other Priorities

Does your team have other priorities or concerns for future data analysis tool development?

- General concerns about data formatting
- Need to work with the calibration pipeline
- Reiterated concerns raised elsewhere in the survey

Solar system geometry

Body coordinates:

Observations of solar system objects need to be mapped in body coordinates, such as planetary latitude and longitude.

Need to account for proper motion and rotation.

Systems like APT already support these coordinate systems.

Detailed and precise pointing information:

If telescope pointing is precisely known, then body coordinates can be determined automatically.

But often there are pointing errors that necessitate corrections using the planetary limb, etc. (e.g., with HST/WFC3).

This will be a challenge for instrument FOVs entirely on the planet's disk.

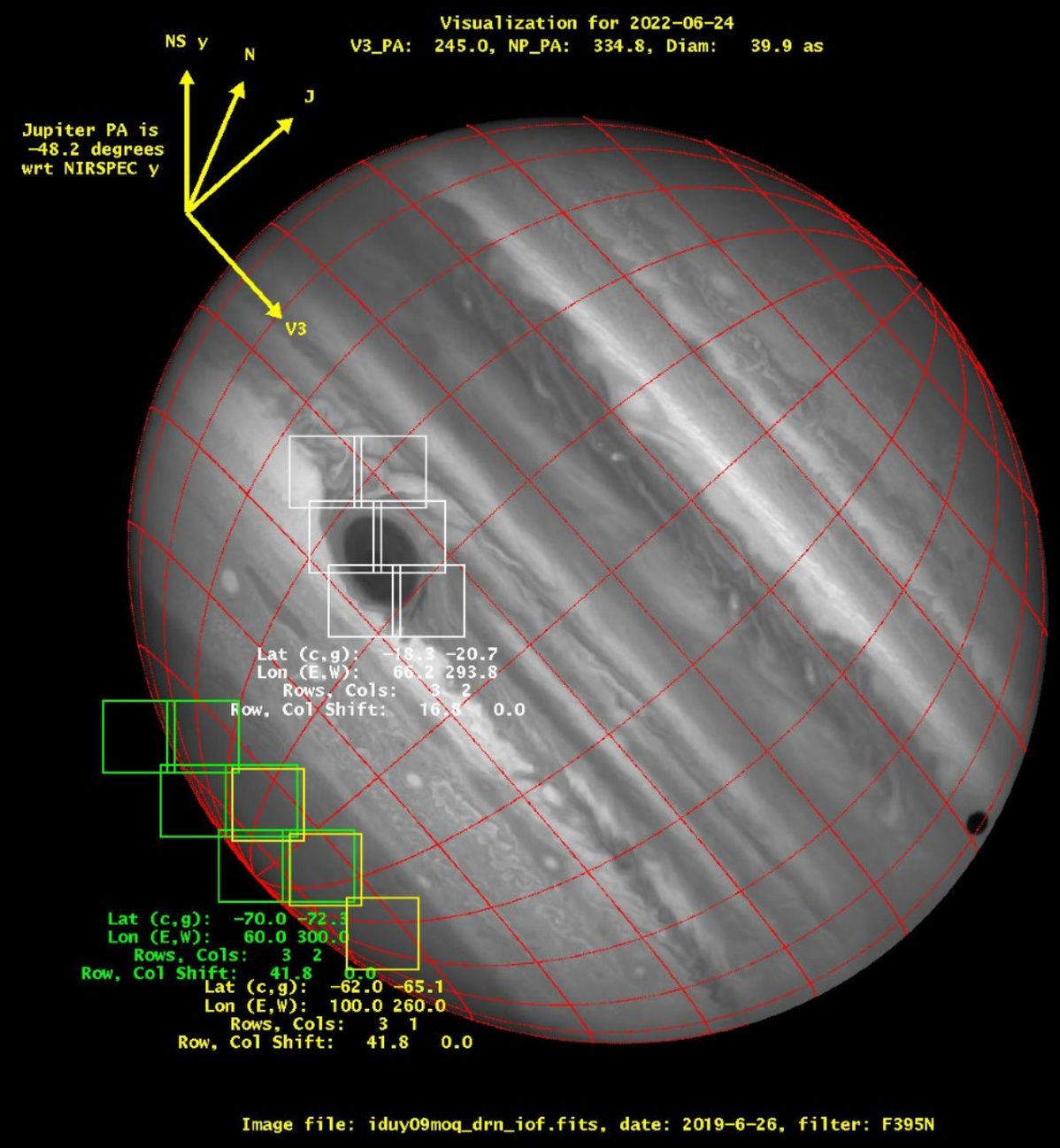


Image: JWST ERS-1373, NIRSpec mosaic layout by Pat Fry (U. Wisc.)

Solar system geometry

Backplanes:

Provide geometrical angles/coordinates for every pixel or spaxel in the data frame.

Most useful approach for end-user science analysis.

Solar system community efforts:

- GTO/ERS team members may hold an online mini-workshop to discuss this issue in Jan. 2021.
- The NASA PDS Rings Node team is developing the OPUS search tool to access solar system observations from the MAST archive, and has plans to add backplanes. This could conceivably be extended to JWST. [link]

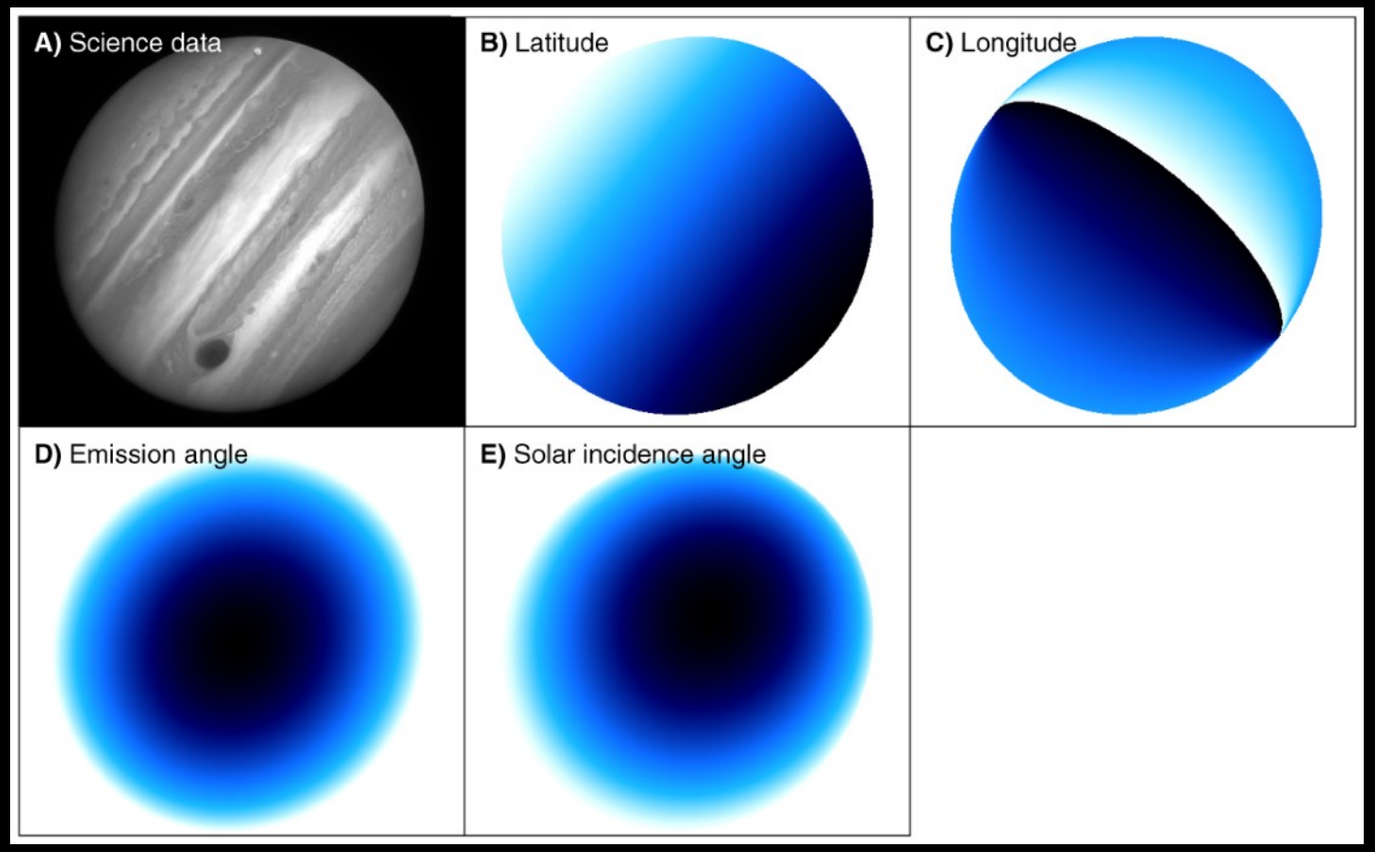


Image: Backplanes of latitude, longitude, solar emission angle, and solar incidence angle for a frame of HST/WFC3 data. From Fig. 5 of Wong et al. (2020). [link]

Takeaway: The Solar System community is concerned about efficiently obtaining precise geometric solutions for SS targets. Relevant information for computing backplanes must be preserved throughout the DRP.

Recommendations: General Principles

- Users expect to analyze their JWST data using a wide variety of tools.
- The best way to support rapid exploitation of JWST data is therefore to foster an ecosystem of new and existing tools that is reliable and robust.
- Goals:
  - Promote new general purpose tools that address the unique characteristics of JWST.
  - Remove roadblocks that inhibit the implementation of existing and custom tools.

Recommendations

1. Maintain emphasis on tools that produce analysis-ready data products rather than top-level results.
   - Prioritize working with instrument-specific data (simulated for now)
     - E.g., spectral extraction (rather than spectral modeling), multi-band photometry (rather than SED fitting, photo-z's, etc.)
   - Suggestions from DA tools survey will be communicated to DAT team
2. Maximize the impact of existing and custom tools
   - Create ASDF interpreters for non-Python languages
     - C/C++ (in progress?)
     - IDL
     - Others?
   - "Stress test" high-value legacy tools on simulated data
     - Run test cases for JWST compatibility
       - ds9 (in progress)
       - QFitsView
       - SExtractor
       - DOLPHOT? Others? Get input from ERS teams.
     - Collaborate with community groups (ERS/GTO)
     - Communicate lessons learned via JDox

Recommendations

3. Enhance support for use of intermediate DRP products
   - Resolve any remaining uncertainties with data formats, header information, etc.
   - Maintain careful documentation of data produces at all levels
   - Highlight notebooks that access lower-level data produces
   - Create non-Python ASDF interpreters
4. Evaluate/implement requirements for mapping Solar System objects
   - Preserve (and document) information needed for backplane reconstruction throughout DRP
   - Consult with SS ERS/GTO teams about requirements