
STScI | SPACE TELESCOPE SCIENCE INSTITUTE
EXPANDING THE FRONTIERS OF SPACE ASTRONOMY

Interactive Analysis Platform

At STScI, a Science Platform is a deployment of the JupyterLab environment on Amazon Web Services, providing users with the necessary software, storage, and compute resources to work with data in the Cloud. **A successful Science Platform should:**

- **Provide a stable software environment** to ensure reliable access and performance.
- **Help users understand how to work within a Cloud framework**, facilitating a smoother transition and efficient use of cloud capabilities.
- **Accelerate science and open new ways of doing science** by fostering innovation through accessible and scalable resources.
- **Be collaborative in design and serve as a gathering space for communities**, promoting interaction and shared learning.

The Science Platforms is not intended to:

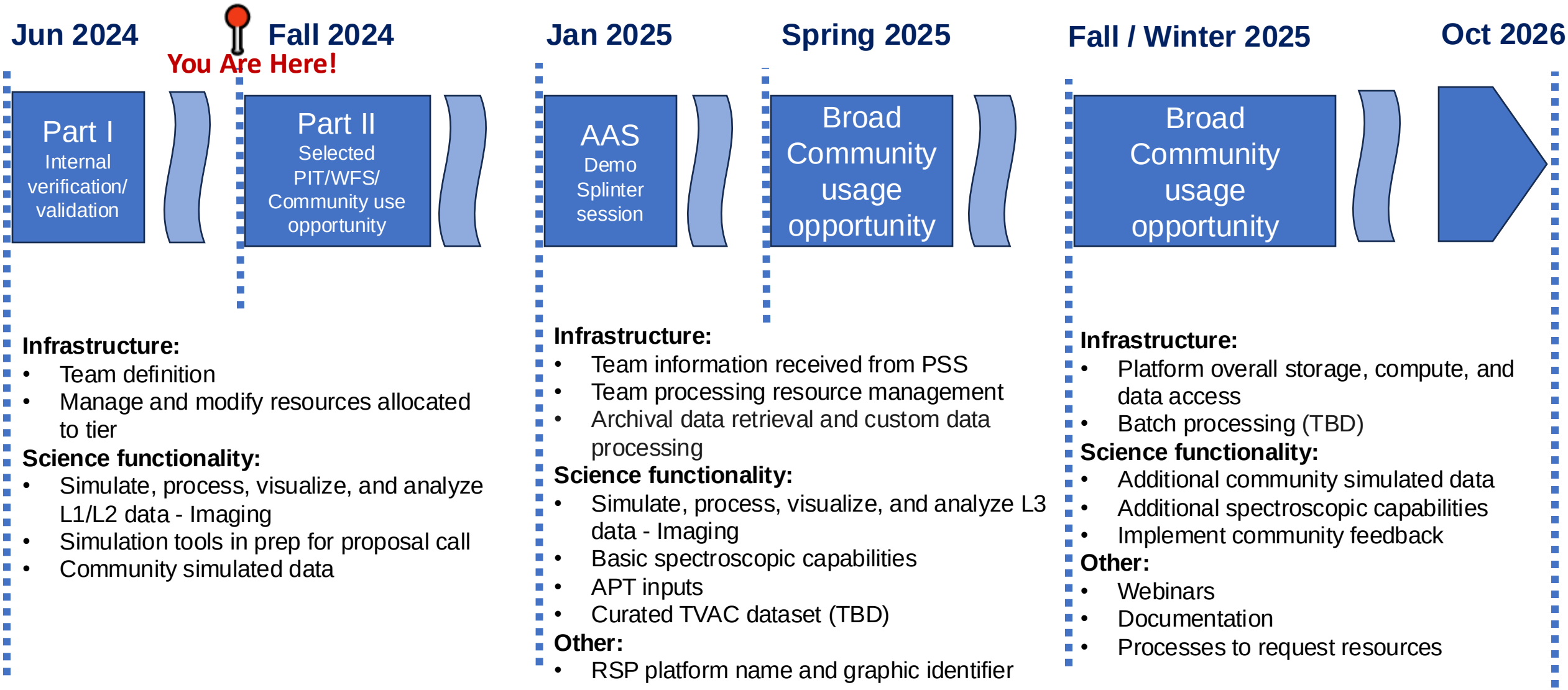
- Replace a high-performance computing center.
- Serve as a one-size-fits-all solution that everyone will use.
- Solve all the challenges associated with working with Big Data.

Science Platform Mission Statement

The Roman Science Platform will enable the community to perform transformational science on Big Data by providing a rich computing environment that will allow broad, low-barrier access to data, compute, and software resources.

Work to meet this vision is organized into **three pillars**:

1. **Collaboration within groups**: file sharing; shared disk space for private teams; open shared disk space; shared CPU resources for teams who request it, etc.
2. **Community support**: documentation; training/workshops; Jupyter Notebook Tutorials; helpdesk, etc.
3. **Software, Data, and Compute Systems**: SOC- and SSC-supported software; access to Roman data (++); software/tools contributed by YOU.


Jun 2024
Fall 2024
You Are Here!
Jan 2025
Spring 2025
Fall / Winter 2025
Oct 2026
Part I
Internal verification/ validation
Part II
Selected PIT/WFS/ Community use opportunity
AAS
Demo Splinter session
Broad Community usage opportunity
Broad Community usage opportunity
Infrastructure:
• Team definition
• Manage and modify resources allocated to tier
Science functionality:
• Simulate, process, visualize, and analyze L1/L2 data - Imaging
• Simulation tools in prep for proposal call
• Community simulated data
Infrastructure:
• Team information received from PSS
• Team processing resource management
• Archival data retrieval and custom data processing
Science functionality:
• Simulate, process, visualize, and analyze L3 data - Imaging
• Basic spectroscopic capabilities
• APT inputs
• Curated TVAC dataset (TBD)
Other:
• RSP platform name and graphic identifier
Infrastructure:
• Platform overall storage, compute, and data access
• Batch processing (TBD)
Science functionality:
• Additional community simulated data
• Additional spectroscopic capabilities
• Implement community feedback
Other:
• Webinars
• Documentation
• Processes to request resources
Phase I
Phase II
Phase III
Roman Science Platform

A Collaborative Hub

- **Team Organization:** Users can be organized into pre-arranged teams.
- **File Access within Teams:** Users have the ability to access files from any team they belong to, regardless of which team they log into the platform with.
- **Computing Resource Selection:** Users in teams can choose from a variety of computing resources, such as CPU and RAM specifications.
- **Usage Visualization:** Users can aggregate and visualize their platform usage metrics (CPU, RAM, egress, storage) over time.
- **Persistent Storage:** Each user is provided with persistent file storage.

Coming Up: Integration of “Real-time collaboration” servers.

Science Functionality

Organization of Science Functionalities in Tutorials and Science Workflows:

- **Tutorial:** A Jupyter notebook that demonstrates how to use specific codes or tools.
- **Science Workflow:** A combination of documentation and notebook tutorials aimed at accomplishing a science-focused use case.

Usage: While individual Jupyter Notebook tutorials can be run as standalone tools, Science Workflows are crafted to provide a complete end-to-end experience.

Implemented Workflows for the Roman Wide Field Instrument (WFI):

- WFI Observation Planning
- WFI Data Simulation
- WFI Data Analysis

Science Functionality

WFI Data Simulations and Analysis

- **Access Data:** Retrieve Roman simulated data from an S3 bucket, as well as existing datasets from MAST.
- **File Management:** Open ASDF files using datamodels, access metadata, manipulate arrays.
- **Data Simulation:** Simulate Roman WFI L1/L2 data products and WFI astronomical scenes.
- **Data Processing:** Convert WFI imaging raw data into exposure level products using RomanCal.
- **Visualization and Analysis:** Utilize quicklook capabilities for visualization and basic interactive analysis, perform forced photometry, and measure galaxy shapes.

WFI Observation Planning

- **Exposure Time Calculation:** Use tools to estimate exposure times.
- **Synthetic Photometry:** Estimate the brightness of sources observed with Roman WFI.
- **Simulate Optics:** Generate WFI simulated Point Spread Functions.

Coming up: Spectroscopic Capabilities; L3 Data Simulation/Processing/Analysis; Curated subset of TVAC data

RSP workshop: September 30th – October 1st 2024

Workshop Goals: Conduct hands-on training sessions to gather feedback, **understand users scientific needs, develop new science workflows**, and identify 'early adopters' interested in using the system for precursor science prior to the Roman launch.

- **Jupyter Notebook Tutorials and Science Workflows:** Evaluate the effectiveness and user-friendliness.
- **Documentation:** Assess clarity, comprehensiveness, and accessibility.
- **User Experience:** Gather insights on the ease of use and functionality of the platform.
- **Team Collaboration:** Review the ability to work within teams, such as file sharing capabilities.

RSP workshop: September 30th – October 1st 2024

Happy faces working on the science platform!

Forty participants from major project partners—both in person and virtual.

Two days of work, including self-guided training and open-format testing.

A few initial hurdles, but ultimately an incredibly successful event—this is why we test!

An incredible showcase of the possibilities on the current infrastructure: from the detection of faint companions to spectroscopic simulators, and highlighting the importance of combining Roman and Rubin data to measure photometric redshift.

RSP workshop: September 30th – October 1st 2024

More than one hundred individual feedback comments.

Next steps: distill, collate, and prioritize feedback. Build an action plan to be shared with all stakeholders.

Next training opportunity: AAS workshop on Sunday, January 12th.

Roman APT Feedback

Please record all issues, suggestions, or any other feedback in the table below. You may add screenshots to the table as needed. If someone else has already entered your issue, please add your name or initials in the second column to indicate your agreement. Understanding how many people were affected by each issue will help us prioritize future work on the RSP. Additionally, please don't forget to select a category from the list provided below. Feel free to copy and paste the appropriate category.

Category:

RSP DOCUMENTATION | TUTORIAL - CONTENT | TUTORIAL - DOCUMENTATION | INFRASTRUCTURE | TOOL

prism

Reference Number	Name of person writing the feedback	Name/Initials of others who agree	A full description of the issue encountered, suggestion, or other feedback	Category
1	Gisella De Rosa	Ori Fox	I really love the landing page, however I would like more details on the history of the platform name. How was it picked? Why?	RSP DOCUMENTATION
2	Ben Rose	Grant Merz, Yash Ejjagiri	There are currently 3 CPU size classes. Is there going to be a class with access to GPUs?	INFRASTRUCTURE
3	Jonathan Hargis	Russ Ryan	Not clear from the landing page how to open a terminal. Also suggest modifying the "how to install extra software" to explain how to open a new terminal	RSP DOCUMENTATION
4	Ben Rose	Bryce Wedig (I've opened PR #59 with a fix)	The simulation documentation (`simulated_data.md`) does not link to the `data_discovery_and_access.ipynb` tutorial. It currently states "More details on accessing these products can be found HERE." without a link.	TUTORIAL - DOCUMENTATION RSP DOCUMENTATION
5	Harry Ferguson		The "installing software" page seems to be focused only on python. I'm interested in installing SExtractor++ (written in C++). https://github.com/astrorama/SourceXtractorPlusPlus. Update: they have a conda installation method, which at least executed for me. I haven't yet tried to run it... Update – after reinstalling in its own python 3.11 environment (using their conda installation instruction) and revising the 01_basic/run_basic.sh script from their 2022 ADASS presentation I got it to run.	RSP DOCUMENTATION
6	Harry Ferguson		When I spawn a terminal, I end up in ~/references/roman_notebooks/content. Seems like it might make more sense to end up in my home directory.	INFRASTRUCTURE
7	Julie McEnery		In the what-is-rsp, under "Who is the RSP for?" The RSP is intended for users interested in preparing observing programs and/or working with data from the Roman Space Telescope Possibly flip the importance of "working with data" and make it more prominent. Perhaps phrase as "..for users interested in generating and/or exploring Roman simulations and working with Roman data.	RSP DOCUMENTATION
8	Rebekah Hounsell	AC	Add units to catalog output `gaia_catalog` in "How to Generate L1 and L2 WFI Files with Roman I-Sim"	TUTORIAL - CONTENT
9	Susana Deustua	Ben Rose	on the romanisim.ipynt landing page - the text on how to start the Roman Calibration kernel might want to point out that the first time one sees the Python 3 Kernel 'button', clicking on that opens a pop-up window and within that is a drop down menu, one of the options is the Roman Calibration.	TUTORIAL - DOCUMENTATION

10	Rebekah Hounsell		`Need to address warnings in "How to Generate L1 and L2 WFI Files with Roman I-Sim" tutorial: AsdfPackageVersionWarning:`
			`File 'file:///home/rebekahhounsell/crds_cache/references/roman/wfi/roman_wfi_readnoise_0748.asdf'`
			`was created with extension URI 'asdf://astropy.org/core/extensions/core-1.5.0' (from package asdf-astropy==0.5.0), which is not currently installed`
			`warnings.warn(msg, AsdfPackageVersionWarning)`
				TUTORIAL - CONTENT

11	Russell Ryan	RH	JDaviz and IMviz (in particular), is very laggy. It's very hard to use it efficiently	TOOL
12	Ben Rose		In `data_discovery_and_access.ipynb` the use of `fsspec_kwargs={'anon':True}` is not described. This makes it hard to vary how one can get access to the "s3://nasa-irsa-simulations/openuniverse2024/" bucket.	TUTORIAL - CONTENT
13	Adrien Thob	Bryce Wedig	Very small suggestion: in the starting server page of the platform, the default server size is large but perhaps it should be defaulted to small server	INFRASTRUCTURE
14	Grant Merz		Notebooks in the /references/roman_notebooks/content/notebooks/ directory are read-only. It may be nice to include a note about this in the intro or tutorials markdown files	RSP DOCUMENTATION
15	Sylvia Kowalski		After signing into the RSP on a Safari browser and launching a terminal (using the blue plus "new launcher" button → "terminal" button) the default terminal view is Settings → Theme → JuptierLab Lite which has black font on a black background (terminal text cannot be seen). Manually setting the view to either JuptierLab Dark or JupiterLab Dark High Contrast the text is visible. Can we change the default setting to be visible out of the gate? **Addition:** I noticed in slack conversations from the first day of training that Ben Rose was also not able to see the terminal, "Currently I don't get a curser. It loads a fully blank page." and I think Ben and my's problem was similar! Ben was also using Safari when this happened.	INFRASTRUCTURE