



**STScI** | SPACE TELESCOPE  
SCIENCE INSTITUTE

EXPANDING THE FRONTIERS OF SPACE ASTRONOMY

# Meeting the "big data" challenge of the Roman era

---

Claire Murray

AAS 244

*Data Management Division, Space Telescope Science Institute*

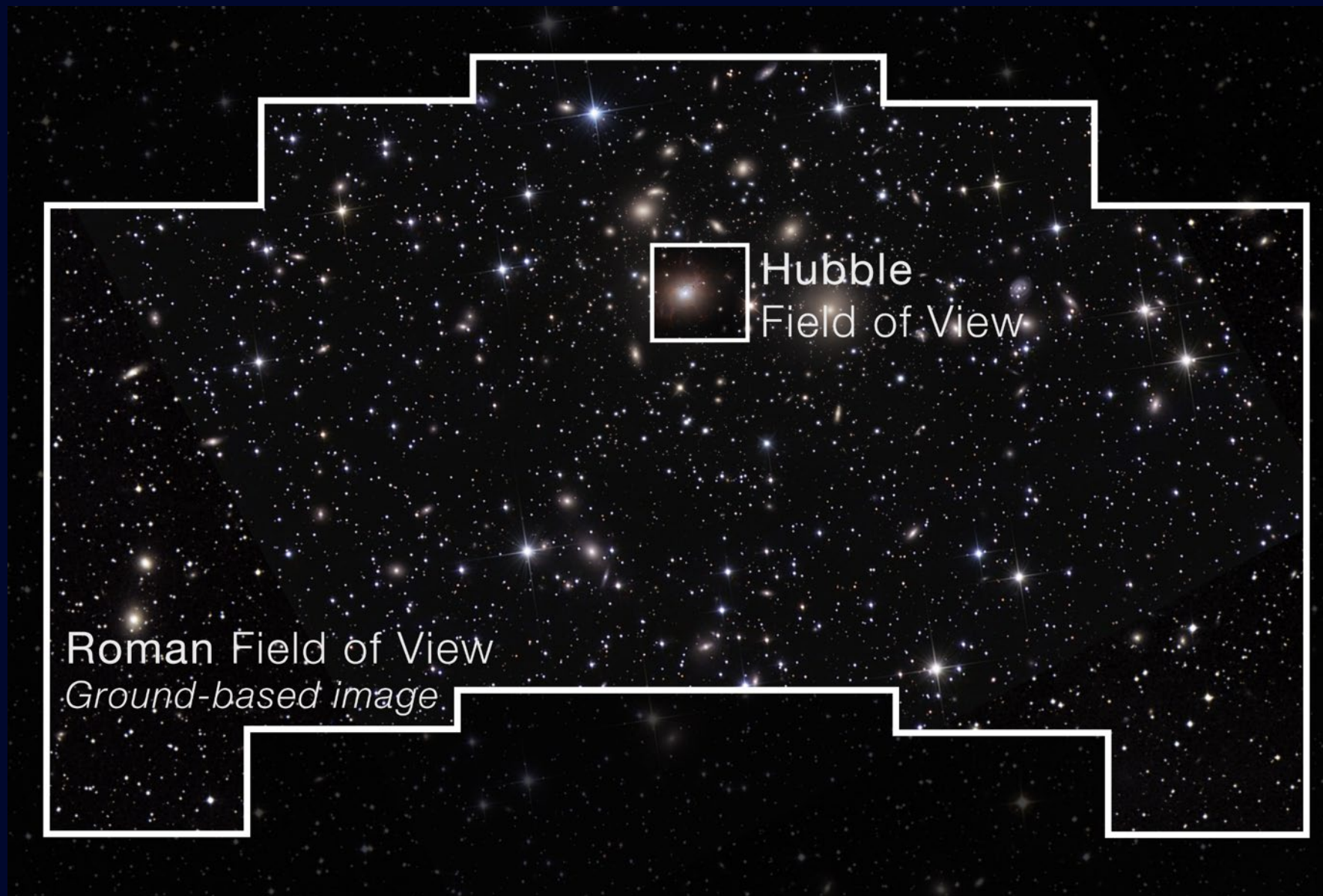
Roman is a *survey* space telescope

*Dark Energy*

*Expansion of the Universe*

*Exoplanets*

*Transformational Astrophysics*



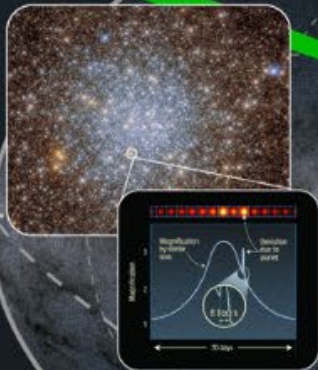
# Roman's Community-Defined Surveys

*Representative footprints only!*

*All Roman data will be public,  
no proprietary period*

e.g.,  $\sim < 15$  min  
cadence  
observations over  
few  $\text{deg}^2$  towards  
Galactic bulge

## Galactic Bulge Time Domain Survey



Early definition General  
Astrophysics survey

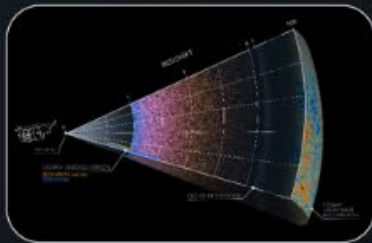
Galactic Plane Survey  
( $\sim 1$  mo)

Galactic Plane

Ecliptic Plane

e.g., wide area (of  
order  $2000 \text{ deg}^2$ )  
multiband survey with  
slitless spectroscopy

## High Latitude Wide Area Survey



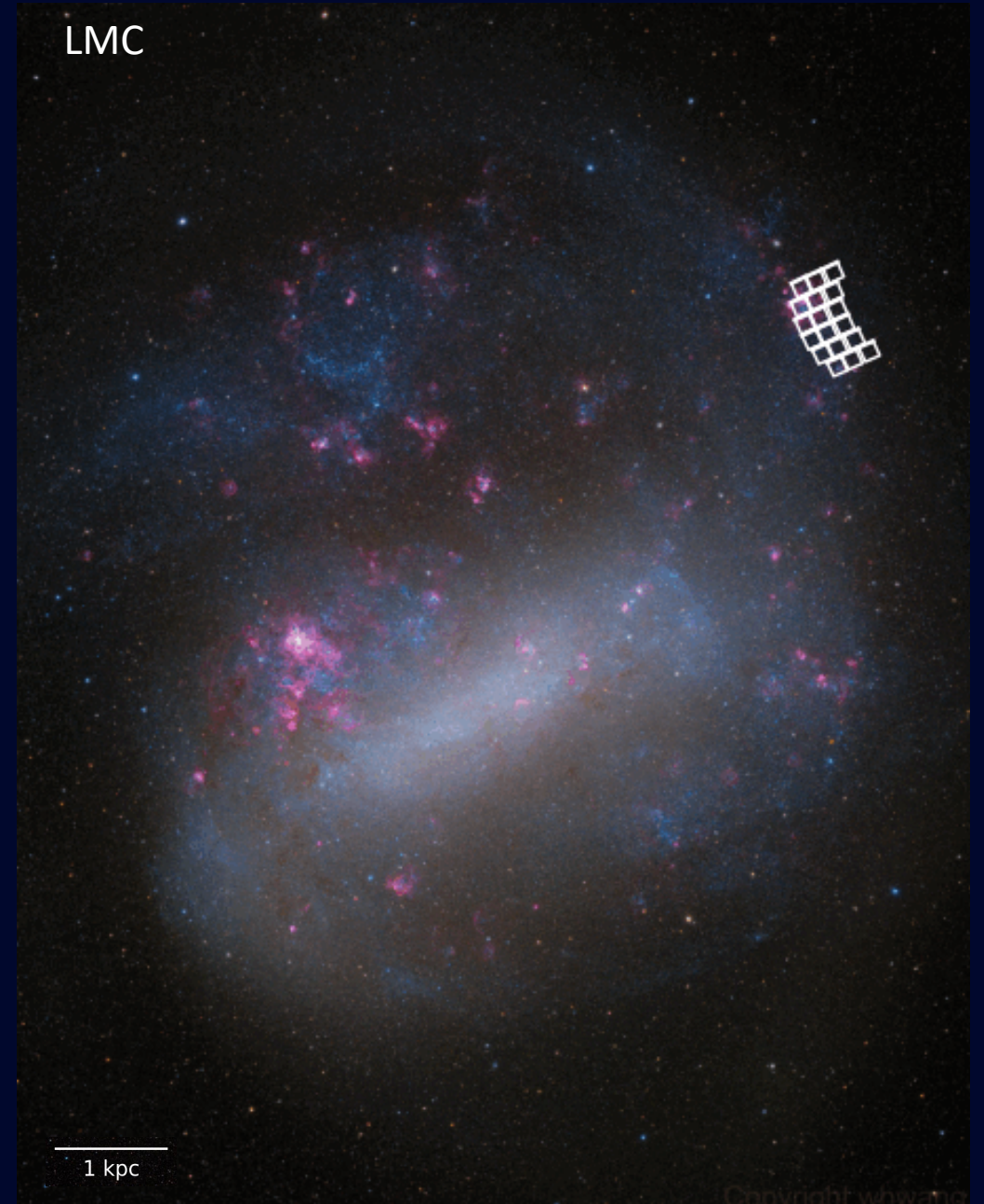
e.g., tiered, multiband time  
domain observations of  
 $\sim 10$ s of  $\text{deg}^2$  at high  
latitudes with slitless  
spectroscopy



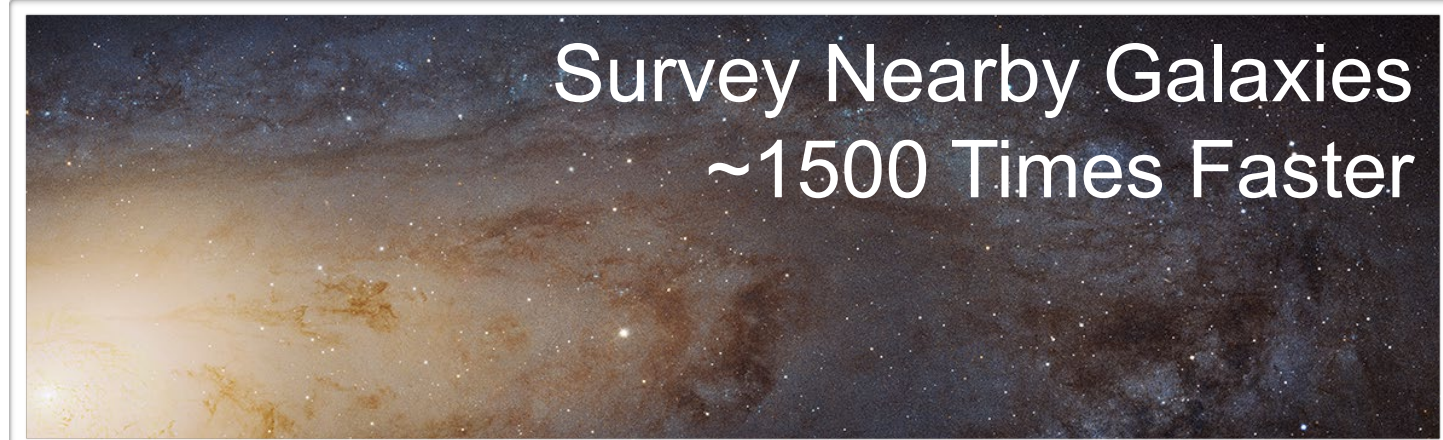
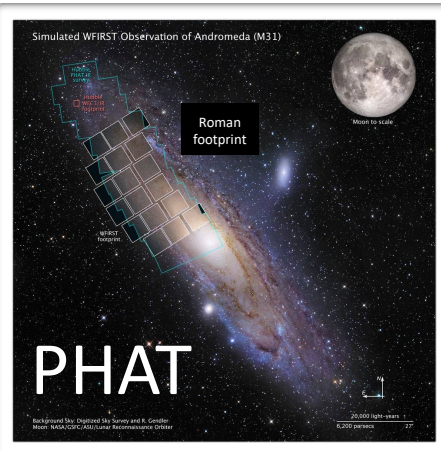
## High Latitude Time Domain Survey

# *The power of a survey telescope*

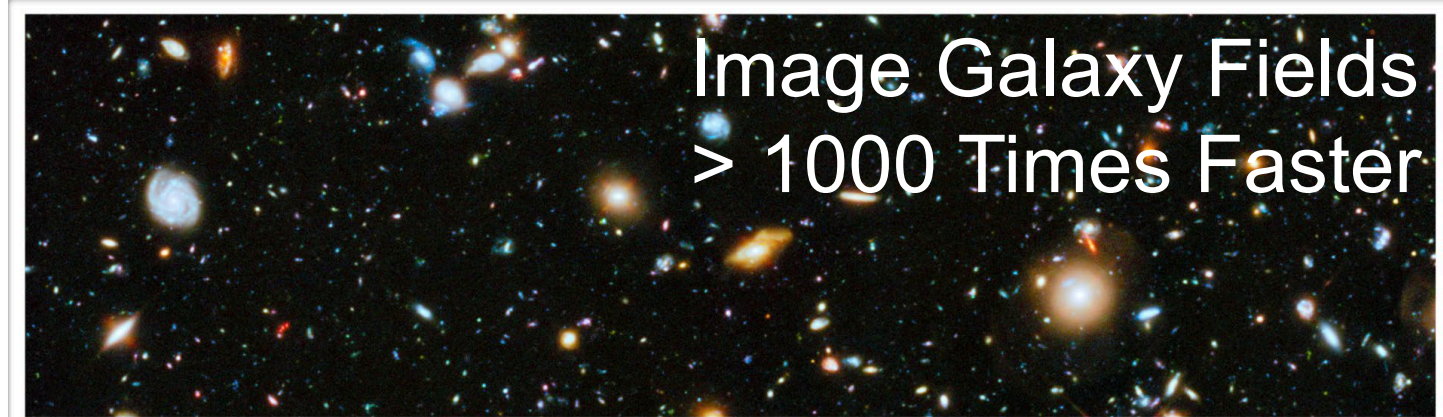
- Speed (survey 1000x's faster)
- Calibration (0.3% uncertainty over 11 mags, 0.5% uncertainty in absolute color)
- Rapid slew & settle times
- No Earth Occultation
- No South Atlantic Anomaly



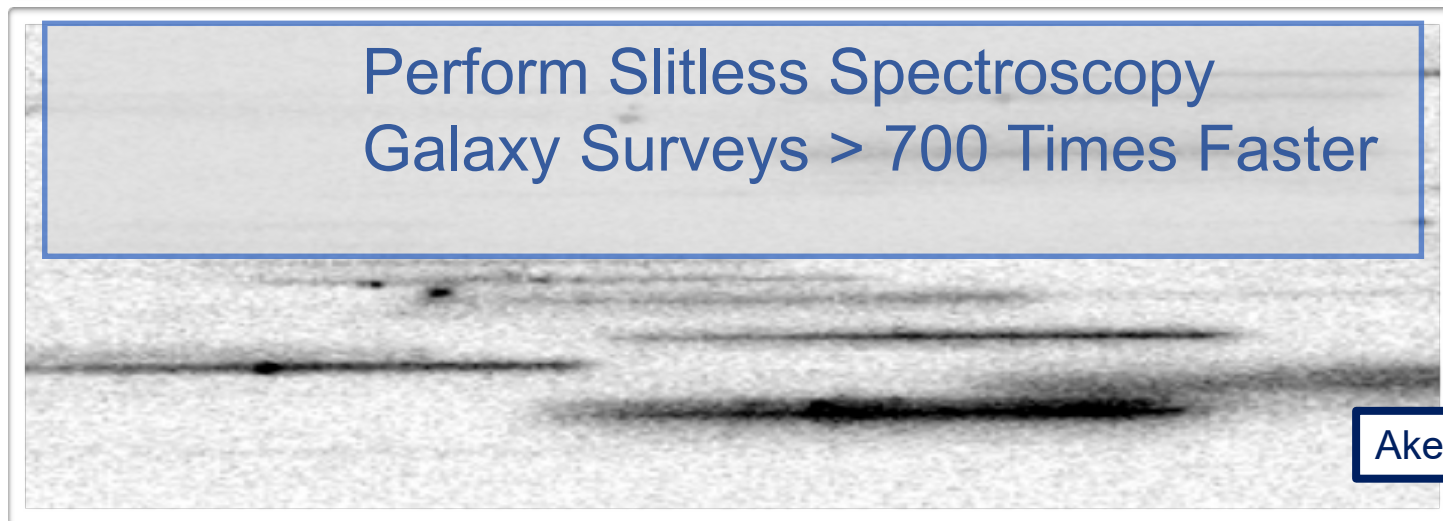
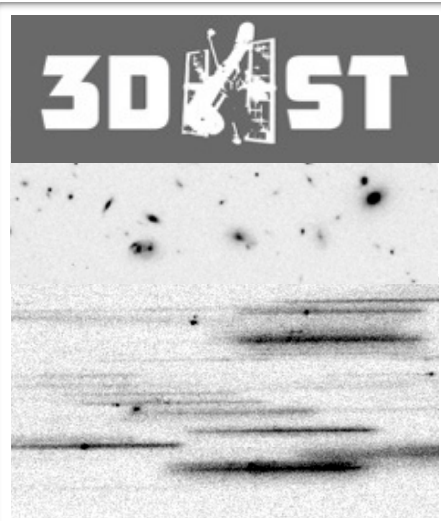
PI J. Dalcanton



PIs S. Faber, H. Ferguson



PI P. van Dokkum



Akeson et al. 2019

# BIG DATA

The Nancy Grace Roman Space Telescope will transmit an unprecedented amount of data from its orbit a million miles away from Earth. Scientists expect it to average almost 1.4 terabytes each day, and after just five years of observations, it should total 30,000 terabytes on MAST, the Mikulski Archive for Space Telescopes. The Hubble Space Telescope sends less than three gigabytes a day, while in the same time even the James Webb Space Telescope will send less than 60 gigabytes.

Hubble

2.7

Webb

58

Roman

1,375

Gigabytes per day sent to Earth



172 terabytes

Hubble's data archive  
1990-2020



1,000 terabytes

Webb's data archive  
after five-year primary  
mission (projected)



30,000 terabytes ( 30 petabytes )

Roman's data archive  
after five-year primary  
mission (projected)



How are we going to help users interact with Roman data?

---



## ***We will bring users to the data*** with the Roman Science Platform

---

A science platform is an integrated set of tools, services, and infrastructure designed to **facilitate and support scientific research and activities**

- Data management and storage
- Computing resources
- Scientific software and tools

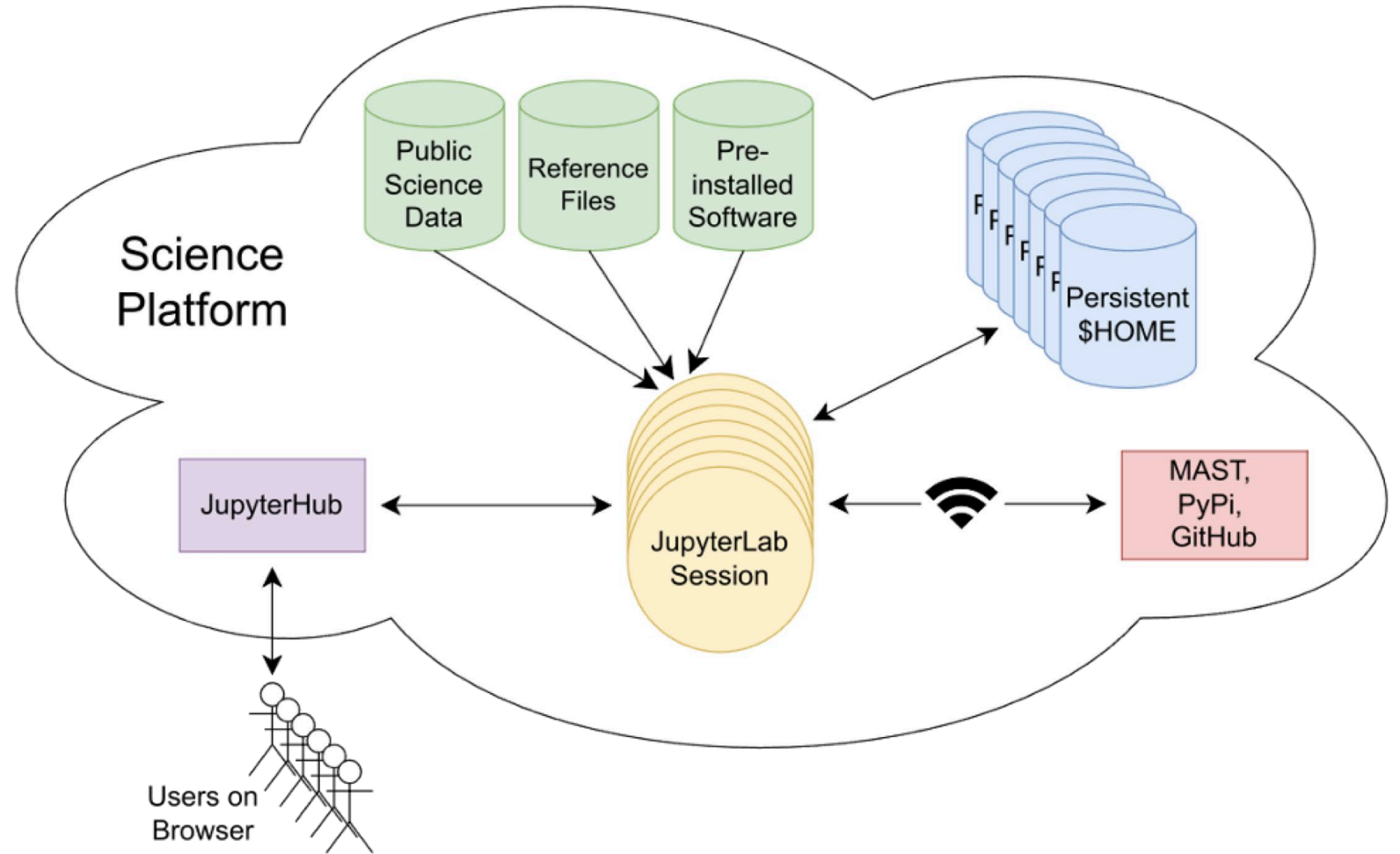




# The structure of a cloud science platform, at a glance

Cloud-hosted environments that enable you to start remote **Jupyterlab** sessions in a browser.

Includes pre-installed software and tools for calibration, analysis, visualization, and training.





# The structure of a cloud science platform, at a glance

Cloud-hosted environments that enable you to start remote **Jupyterlab** sessions in a browser.

Includes pre-installed software and tools for calibration, analysis, visualization, and training.

The screenshot displays the JupyterLab interface. On the left, a file browser shows the directory structure: `/romancal-notebooks / create_science_data /`. A file named `Create_example_level1_wfi_im...` is selected, with a last modified date of `3 months ago`. The main area shows a code editor with the following Python code:

```
[ ]: # make sure that you have these packages installed in your local environment
import numpy as np
import matplotlib
from matplotlib import pylab as plt
import asdf
import astropy
from astropy.time import Time
from astropy.io import fits
import roman_datamodels
import roman_datamodels.stnode as rds

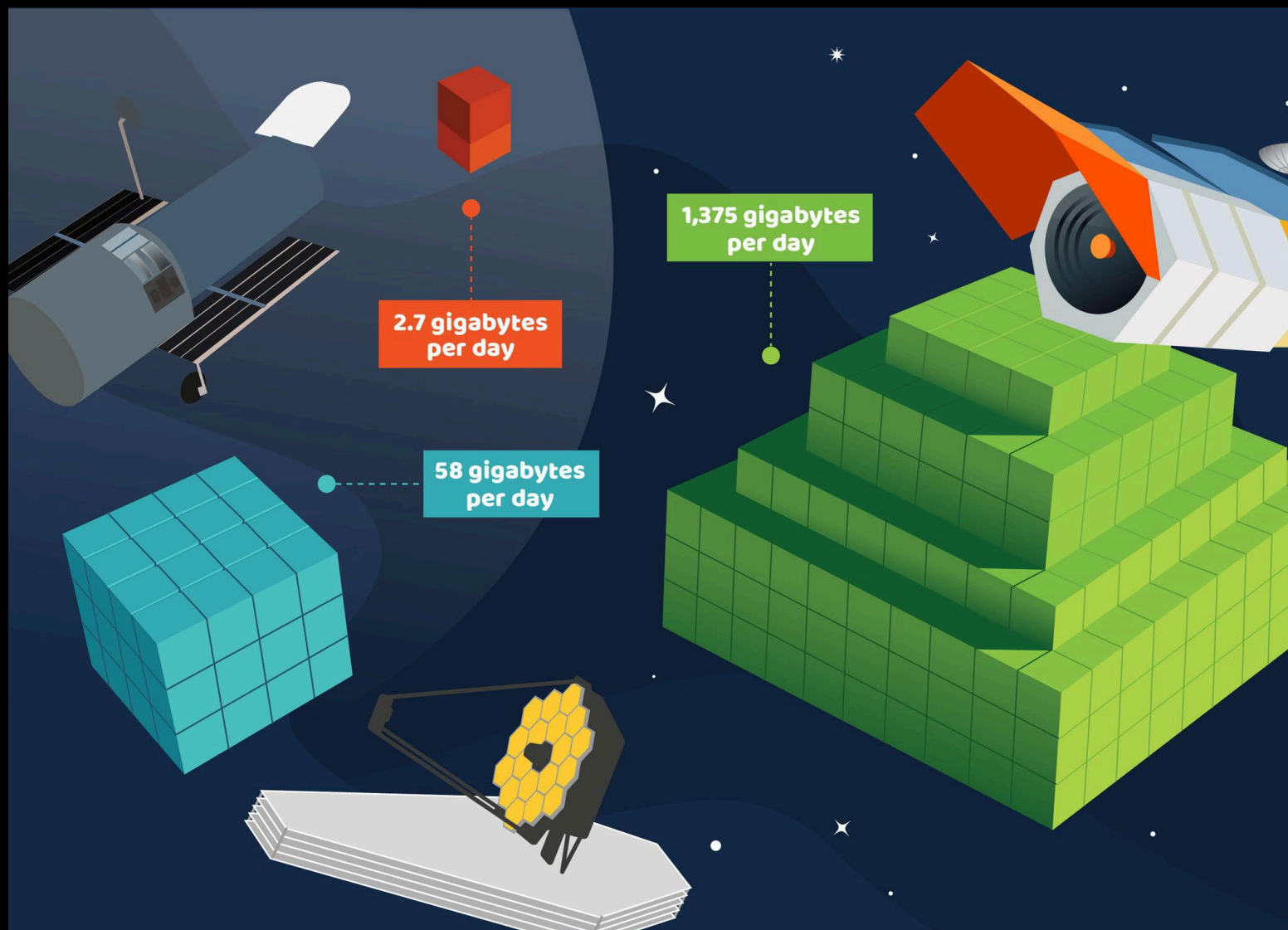
# the following allows us to grab example data from an STScI box folder
from astropy.utils.data import download_file
box_path = 'https://data.science.stsci.edu/redirect/Roman/Roman_Calibration_And_
```

Below the code, a section titled "Files currently available in the box folder linked above include:" lists several filenames:

Filename
r0000101001001001001_01101_0001_WFI01_
r0000201001001001001_01101_0001_WFI01_
r0000201001001001002_01101_0001_WFI10_
r0000201001001001003_01101_0001_WFI08_
r0000101001001001001_01101_0001_WFI01_



# Roman has special science platform needs...





## Vision for the *Roman* Science Platform

---

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.

*Enable collaboration*

*Support the community*

*Build software and tools that bring people to the platform*



# Vision for the *Roman* Science Platform

## *Enable collaboration*

- Shared resources within groups
  - Real-time collaboration tools
- Variety of available computing resources

## *Support the community*

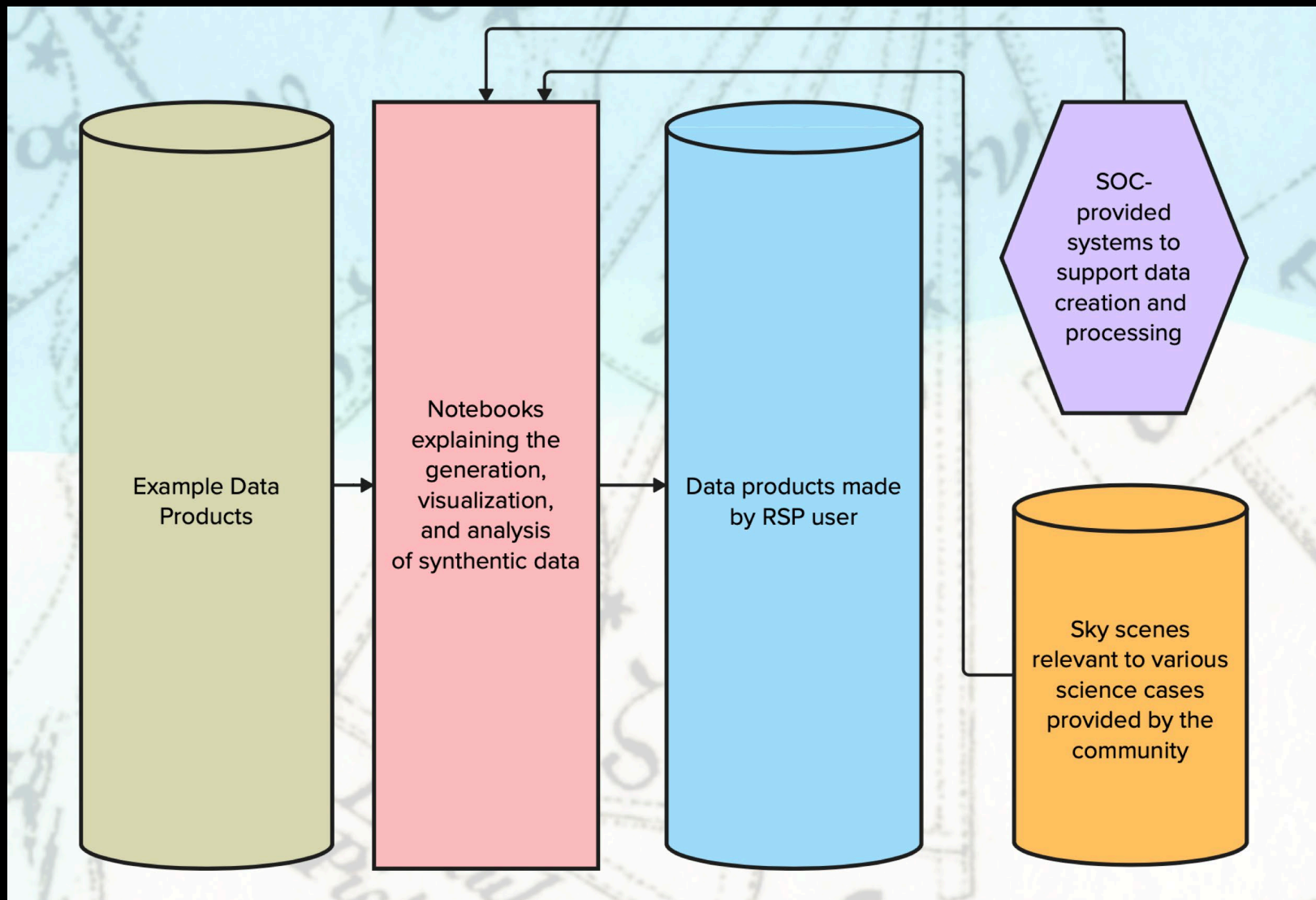
- Regular opportunities for community feedback
  - First round* (STScI+PITs): Summer 2024
  - Second round* (public): AAS 245 (Jan 2025)
- Tutorials, documentation, workshops, helpdesk

## *Bring people in*

- Integrated data stack allows community to interact with Roman tools (pipeline, simulator, APT, visualization and analysis)



# Roman science platform **PRE-LAUNCH** will be a gathering place





# Join the Roman Science Forum today!

**Now:** Hear about mission updates & join working groups

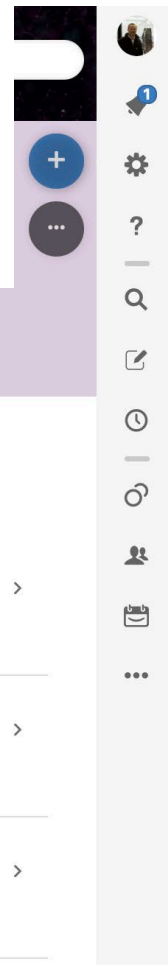
**Coming soon:** training opportunities for the Roman APT, preview the Roman Science Platform

[Mission Committees](#)
[Code of Conduct](#)
[Working Group Signup Form](#)

## Welcome to the Roman Forum

Collaboration focused on maximizing the science achieved with the Nancy Grace Roman Space Telescope's Wide Area and Survey Science. Feel free to browse or, if you'd like to get more involved, join one of the working groups!

[Join a Working Group](#)



### My Confluence Spaces

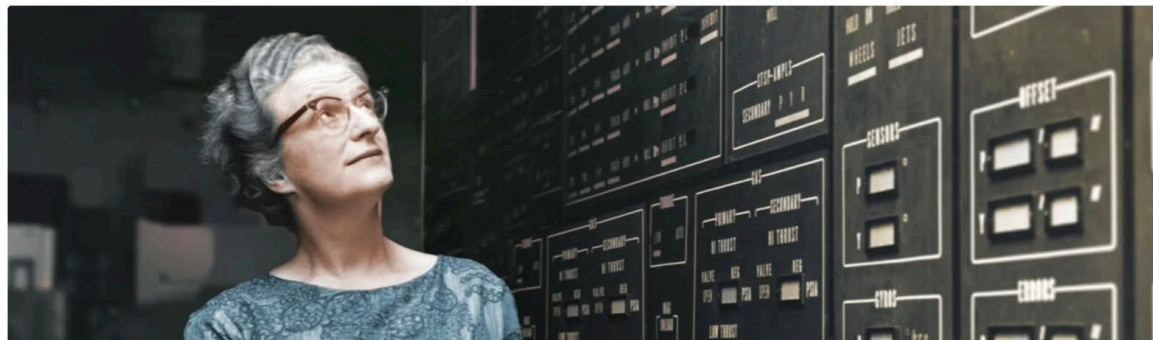
All ▾

#### Working Groups

Working Groups Landing Page

- Community Discussion ★
- Implementation Science WG ★
- WFI Working Groups
  - Calibration WG ★
  - Simulations WG ★
  - Software WG ★

#### Survey Working Groups






This is the home page for Confluence for the Roman Space Telescope at STScI.

**Usage of this system is subject to Roman's [Code of Conduct](#).**

**Setting up an account:** Limited public content will be available on this webpage, including updates on the Roman mission and details on upcoming opportunities to engage with Roman. To access additional content and functionality (such as commenting on posts and interacting

### Latest News

- 
[Roman Summer Conference Overview and Goals](#)
3 hours ago
- 
[Registration Open for Summer Roman Conference](#)
3 hours ago
- 
[This is the Implementation Science WG Blog](#)
3/8/2024



## More ways to get involved with Roman!

- Survey planning underway
  - Galactic Plane General Astrophysics Survey definition committee being formed
  - Galactic Plane Survey – committee is being formed (see newsletter!)
  - Roman Time Allocation Committee (RoTAC): Call for (self-) nominations coming soon (see newsletter!)
- **Roman conference** at Caltech, July 9-12, 2024
  - in-person registration due June 24, virtual registration due June 25



Roman Science  
Conference



Latest Roman  
News!



Join the  
Science Forum!