



STScI | SPACE TELESCOPE
SCIENCE INSTITUTE

EXPANDING THE FRONTIERS OF SPACE ASTRONOMY

The Roman Research Nexus: A Science Platform for Roman Data Analysis

Gisella De Rosa
on behalf of the Roman Science Operations Center (SOC)

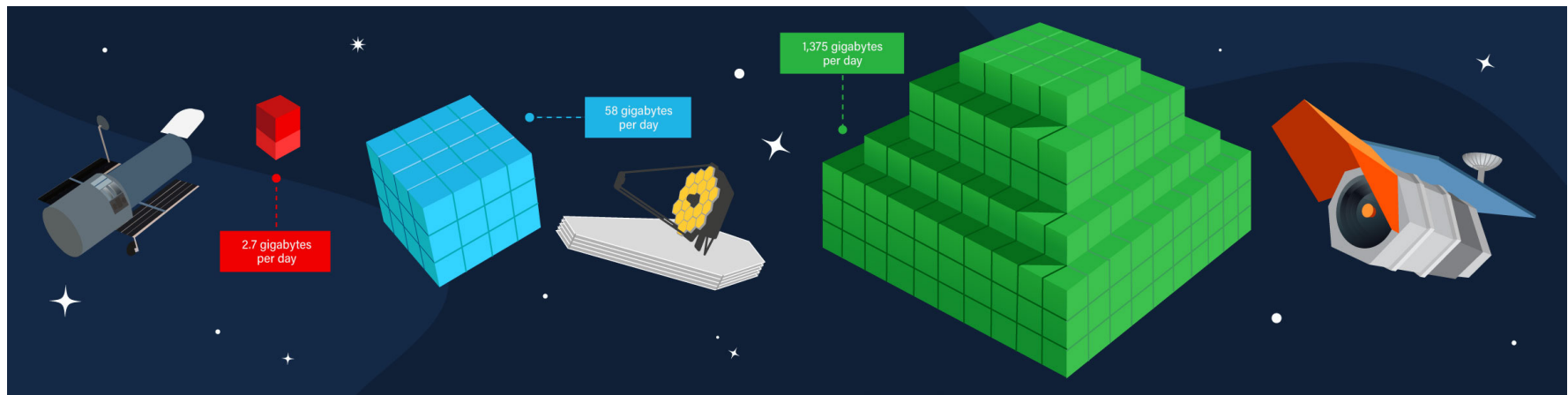
Hot Wiring the Transient Universe – March 25, 2026



Roman changes how we do science with NASA flagship missions



- ~1.4 TB/day of compressed data
- Orders of magnitude larger than Hubble or JWST
- All the data public immediately (no proprietary time)





Roman Requires a Different Analysis Model



- Traditional download-and-analyze workflows do not scale
- Immediate public access requires scalable, shared infrastructure
- Data, compute, and tools must be integrated



What is the Roman Research Nexus



The Roman Research Nexus is a **cloud-based science platform** designed for Roman users.

- Brings compute and tools to cloud-hosted data
- Provides a pre-configured environment with Roman analysis tools
- Supports end-to-end workflows from simulation to analysis
- Provides a shared environment for collaborative science

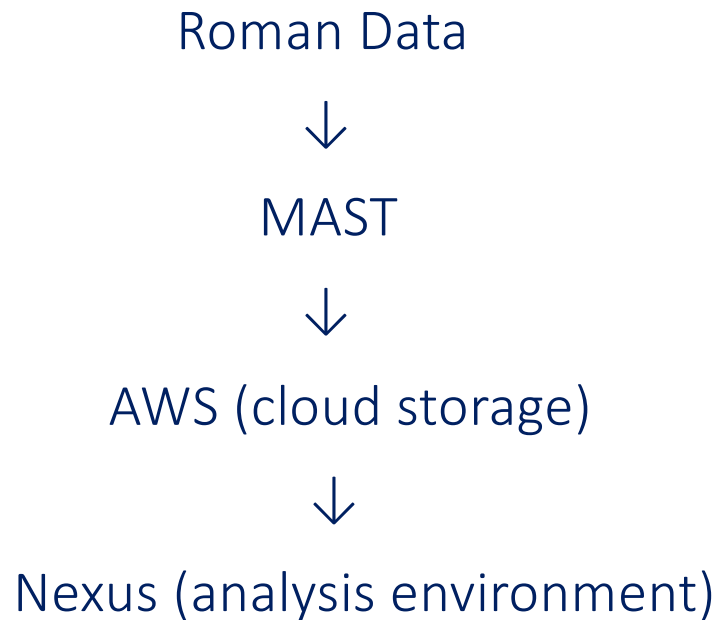


Nexus

Access is available today through MyST accounts, with both individual and team-based environments



Where the Data Live



- Roman data are archived and accessed through MAST
- Data are hosted in AWS East-1
- Includes SOC/SSC & PIT products (e.g. RAPID products)
- The Nexus operates in the same cloud environment → **Data can be streamed, no need to egress (download)**



From alerts to analysis: where the Nexus fits



- RAPID enables alert generation and distribution
- MAST provides access to Roman data products
- The Nexus is where analysis workflows are executed

RAPID alerts → Brokers → Users



MAST



Nexus



A Collaborative Hub



- Team-based environments for Roman science projects
- Shared access to data and resources (compute, storage, egress)
- Real-time collaboration within a common analysis environment
- Supports distributed, multi-institution teams



How to
Team Request

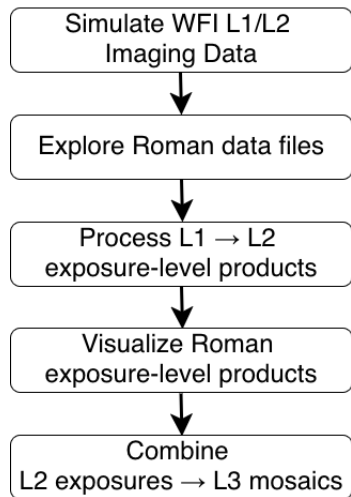


End-to-end scientific workflows

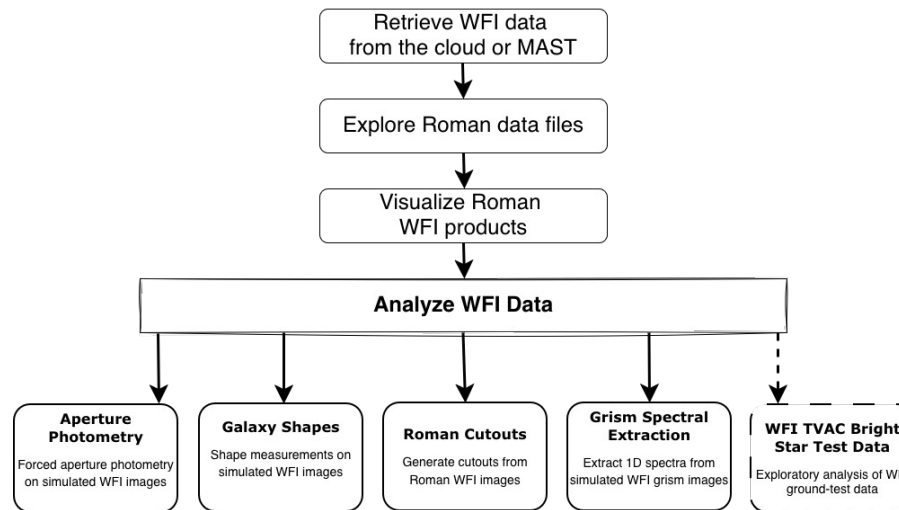


- Simulation → calibration → analysis in a single environment
- Roman tools integrated into workflows (RomanCal, Roman I-sim)
- Supports reproducible, iterative analysis
- Enables development and testing of analysis pipelines

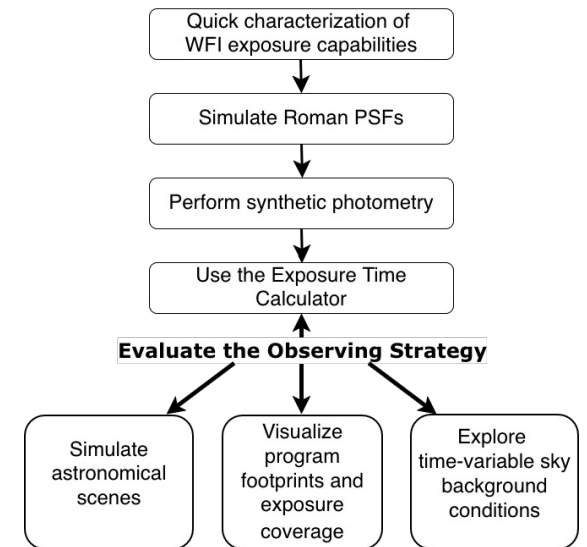
WFI Data Simulations



WFI Data Analysis



WFI Observations Planning





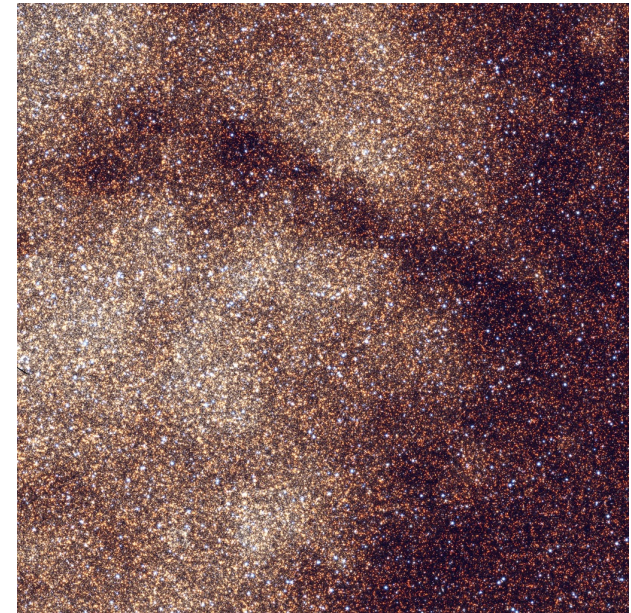
Roman Simulated Data in the Nexus



- Roman simulated data are available directly within the Nexus
- Includes Roman I-sim simulations and OpenUniverse datasets
- Data are provided in pipeline-compatible formats
- Enables testing of workflows and analysis pipelines

COMING UP - CCS simulations with Roman I-sim:

- GBTDS: ~8 days, five bulge fields + Galactic Center
- HLWAS: ~5 deg² all tiers, approximating first-year depth
- HLTDS: ~5 deg² , full pilot and ~90 days of core component

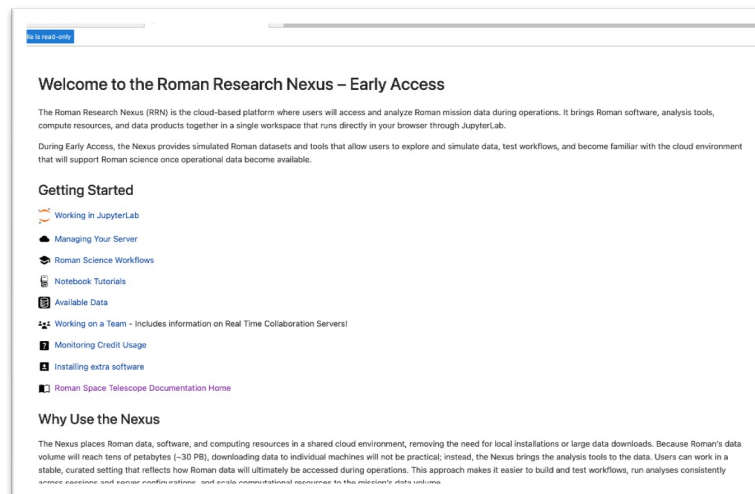




Roman Research Nexus Early Access



- The Roman Research Nexus is available today in Early Access
- Users can access the platform via MyST accounts (individual and team environments)
- Core workflows and simulation datasets are already available
- The platform is actively evolving with new capabilities and datasets





What's Coming on the Nexus (Selected Examples)



New science workflows

- Time-domain simulation and analysis workflows (~April)
- WFI detector characterization workflows (~June)

Compute capabilities

- GPU-enabled compute for ML and large-scale analysis
- Larger server configurations (~100 CPUs) for parallel workflows

Data access and management

- S3 buckets for team-based data storage and sharing
- Bulk upload and download capabilities



Connecting Archive and Analysis

- Roman data are accessed through MAST
- From the Nexus, users can query the archive programmatically (e.g., astroquery)
- We are developing infrastructure to enable workflows that connect archive interfaces and analysis environments
- Goal: a seamless path from data discovery to analysis



Search

Nancy Grace Roman Space Telescope Notebooks

[Roman Research Nexus Data Discovery and Access in the Cloud](#)

[How to Open Roman Data Files \(ASDF\)](#)

Simulating WFI Imaging Data with Roman I-Sim

Calibrating WFI Exposures with RomanCal

Diving Into WFI TVAC Bright Star Test Data

Aperture Photometry

Measuring Galaxy Shapes

Roman Spectroscopy: Grism Spectral Extraction

The Roman Interactive Sensitivity Tool (RIST)

STIPS Tutorial

STPSF Tutorial

Roman WFI Synthetic Photometry with synphot

Pandemia Tutorial

Visualization of Roman APT products

The Roman Background Visualization Tool (RBVT)

working_with_asdf/working_with_asdf.html



Contents

Accessing MAST Data

In this section, we will go through the steps to retrieve archived MAST data from the cloud including how to query the archive and stream the files directly from the cloud, as well as download them locally.

Enabling Cloud Access

The most important step for accessing data from the cloud is to enable `astroquery` to retrieve URIs and other relevant cloud information. Even if we are working locally and plan to download the data files (not recommended for Roman data), we need to use this command to copy the file locations.

```
Observations.enable_cloud_dataset()
```

INFO: Using the S3 STScI public dataset [astroquery.mast.cloud]

Querying MAST

Now we are ready to begin our query. This example is rather simple, but it is quick and easy to reproduce. We will be querying HST WFC3/IR data of M85. In practice, the science platform should primarily be used for analyzing and exploring Roman data products. However due to the smaller file sizes, HST WFC3/IR data provides a nice example. The process is identical regardless of which space telescope is used.

In our query, we specify that we want to look at HST data using the F160W filter and WFC3/IR. We also specify the proposal id to easily get the data of interest. Once we get the desired observations, we gather the list of products that go into the observations. We then filter the products to gather all the level 3 science data products associated with a specific project which still leaves us with 60 data products.

```
# query MAST for matching observations
obs = Observations.query_criteria(obs_collection='HST',
                                  filters='F160W',
                                  instrument_name='WFC3/IR',
                                  proposal_id='11368',
                                  dataRights='PUBLIC')

# get the list of products (files)
products = Observations.get_product_list(obs)

# filter the products
filtered = Observations.filter_products(products,
                                         calib_level=[3],
                                         productType='SCIENCE',
                                         dataproduct_type='image',
                                         project=['CALWFS'])

print('Filtered data products:\n', filtered, '\n')

# filter for just one product
single = Observations.filter_products(filtered,
                                       obsID='24797441')

print('Single data product:\n', single, '\n')
```

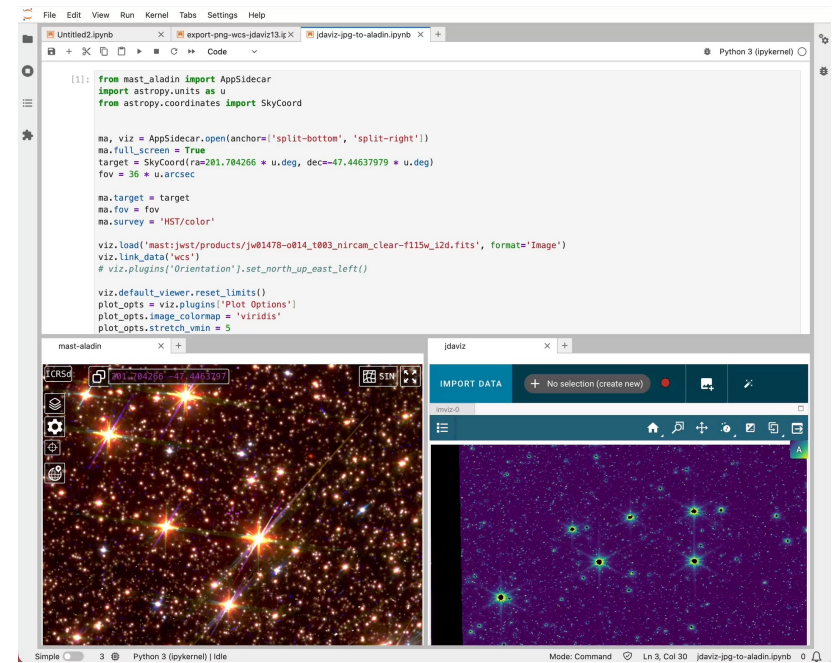
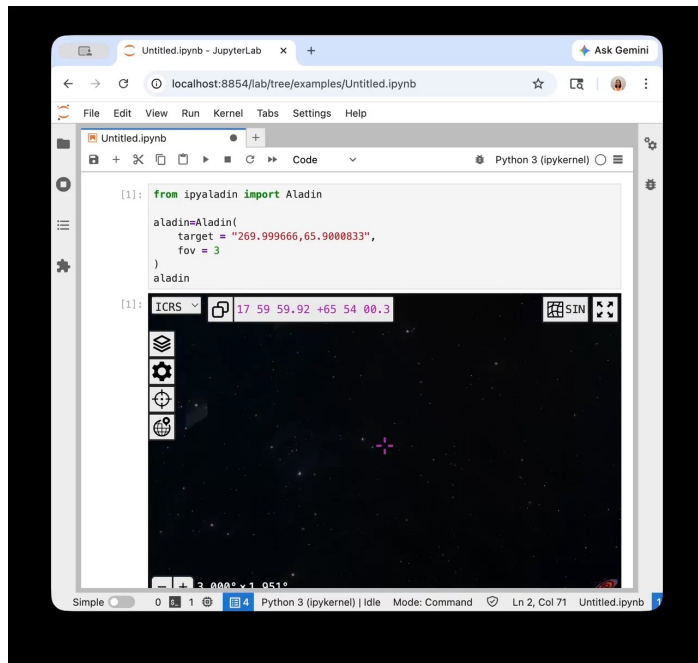
obsID	obs_collection	dataproduct_type	...	dataRights	calib_level	filters
23831959	HST	image	...	PUBLIC	3	F160W
23831959	HST	image	...	PUBLIC	3	F160W
23831961	HST	image	...	PUBLIC	3	F160W



Towards Integrated Data Discovery and Analysis



- New interfaces support interactive data exploration within the archive
- Bridge data discovery and downstream analysis workflows
- Part of broader efforts to integrate archive and analysis environments





External data and co-location



- Co-location is critical for efficient analysis at Roman scale
- Many science cases (e.g., transients) require multi-wavelength data
- External datasets are generally not co-located with Roman data

MAST Ongoing work:

- Develop a cloud-hosted staging area and upload interface for contributors
- Identify key datasets that augment Roman science and ensure they are readily accessible within MAST workflows



<https://www.surveymonkey.com/r/SXC2PCJ>



From RAPID alerts to data access and analysis

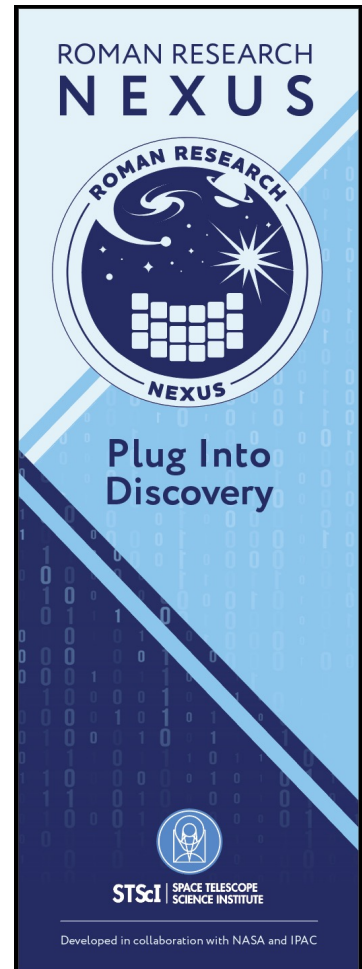
- Access Roman data products and move into analysis
- Develop, test, iterate on your workflows
- Work in shared environments with your collaborators

The Nexus is available now in Early Access (MyST accounts)
Use it, and help shape its development



roman.science.stsci.edu

Plug in. Explore. Discover.



Questions?

Help Desk

