STScI | SPACE TELESCOPE SCIENCE INSTITUTE

EXPANDING THE FRONTIERS OF SPACE ASTRONOMY

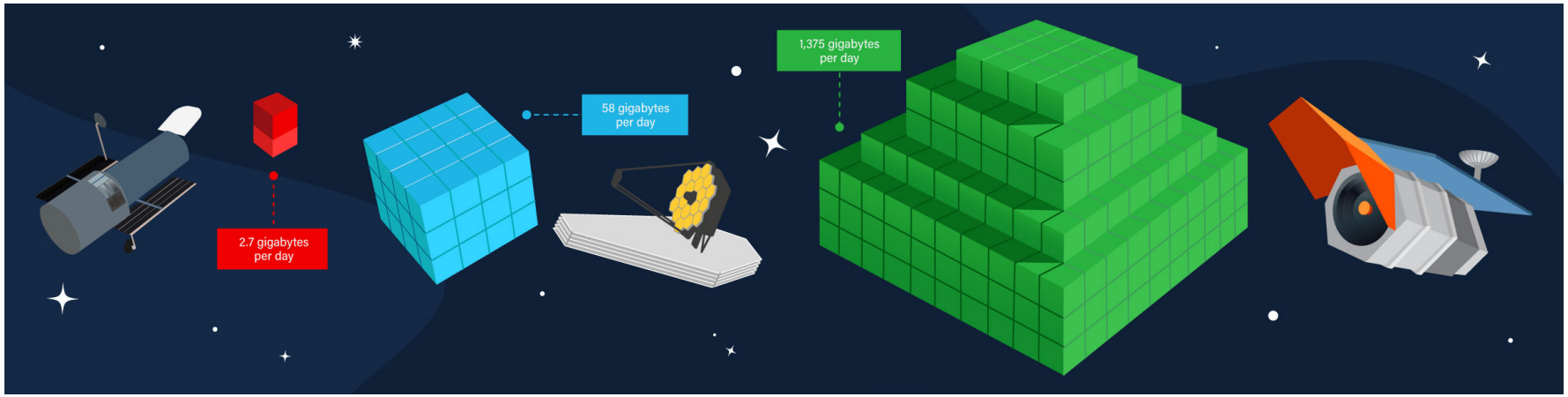
The Roman Research Nexus: A Science Platform for Roman Data Analysis

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

↓

MAST

↓

AWS (cloud storage)

↓

Nexus (analysis environment)

- 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

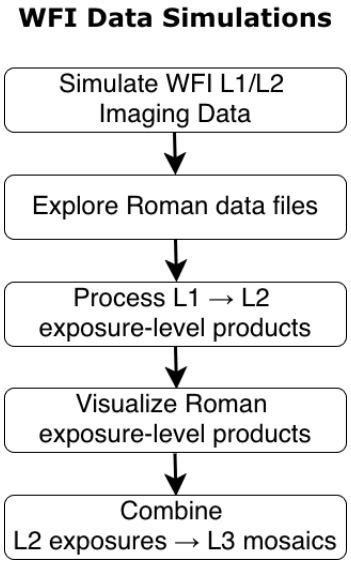
A Collaborative Hub

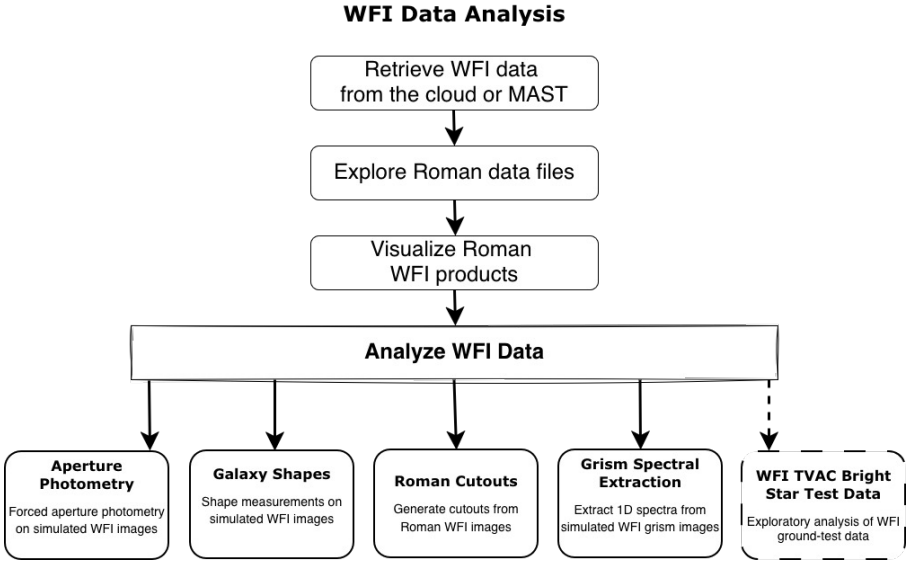
- Team-based environments for Roman science projects
- Shared access to data and resources (compute, storage, egress)
- Real-time collaboration within a common analysis environment
- Supports distributed, multi-institution teams

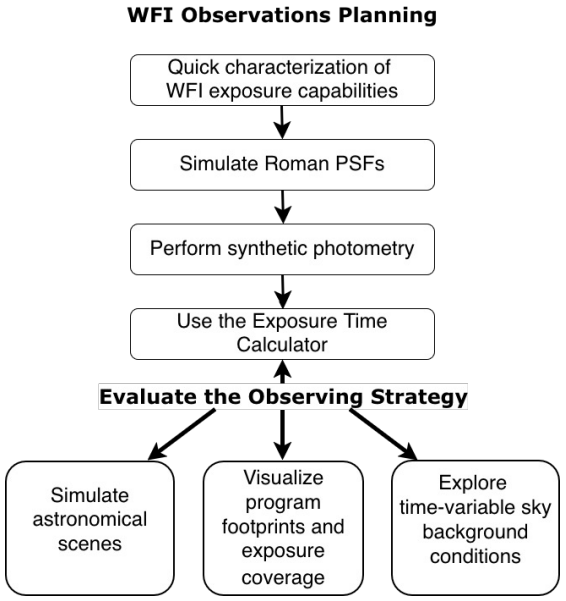
How to Team Request

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





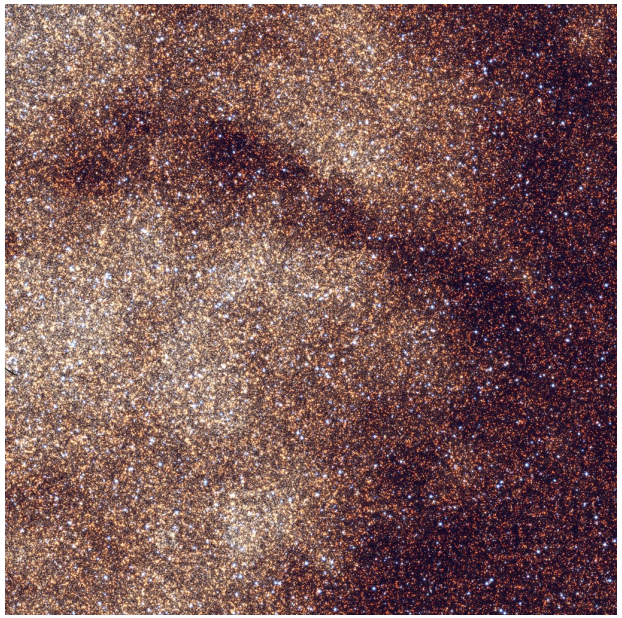


Roman Simulated Data in the Nexus

- 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

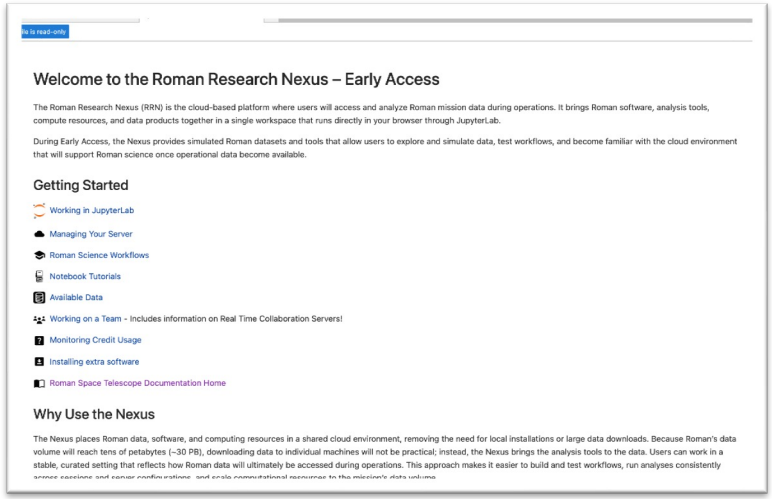
COMING UP - CCS simulations with Roman I-sim:

- GBTDS: ~8 days, five bulge fields + Galactic Center
- HLWAS: ~5 deg^2 all tiers, approximating first-year depth
- HLTDS: ~5 deg^2 , 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

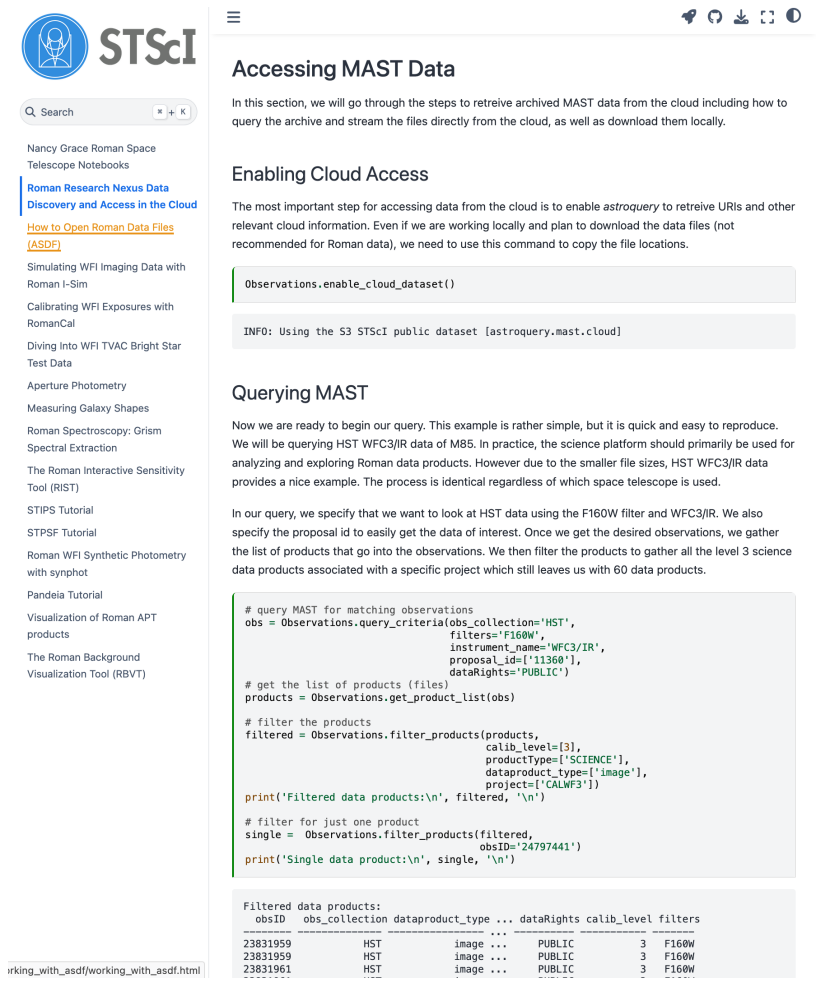
- GPU-enabled compute for ML and large-scale analysis
- Larger server configurations (~100 CPUs) for parallel workflows

Data access and management

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

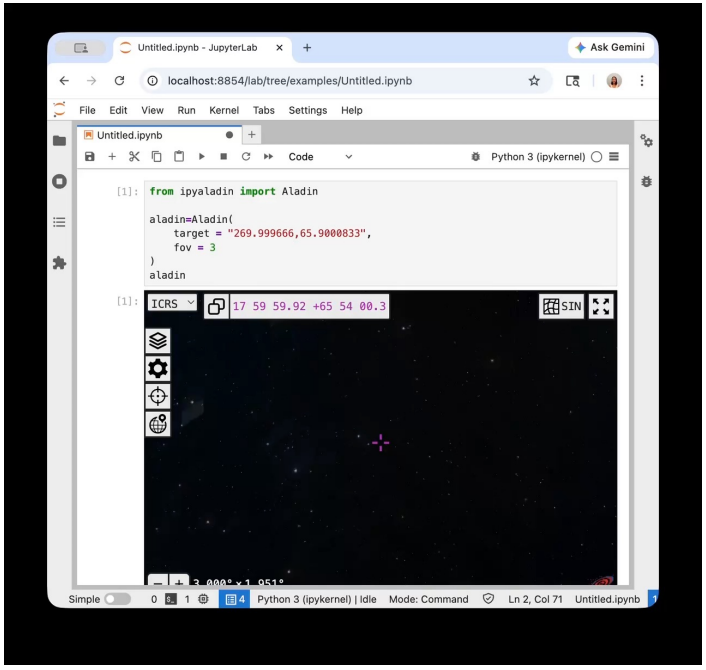
Connecting Archive and Analysis

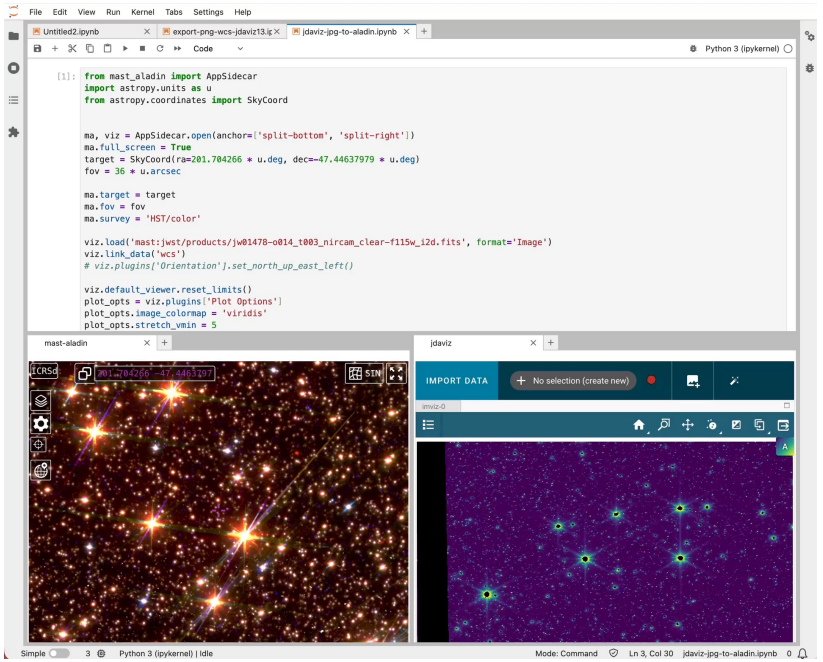
- Roman data are accessed through MAST
- From the Nexus, users can query the archive programmatically (e.g., astroquery)
- We are developing infrastructure to enable workflows that connect archive interfaces and analysis environments
- Goal: a seamless path from data discovery to analysis



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

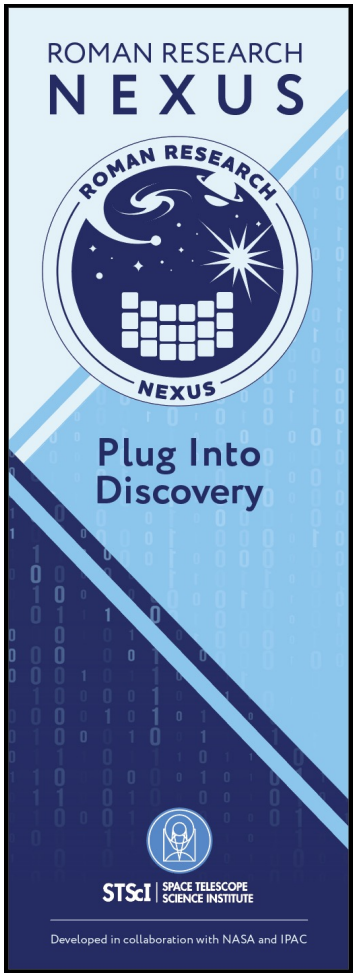
- Access Roman data products and move into analysis
- Develop, test, iterate on your workflows
- Work in shared environments with your collaborators

The Nexus is available now in Early Access (MyST accounts)

Use it, and help shape its development

Plug in. Explore. Discover.

roman.science.stsci.edu



Questions?

Help Desk