



**STScI** | SPACE TELESCOPE  
SCIENCE INSTITUTE

EXPANDING THE FRONTIERS OF SPACE ASTRONOMY

## **Roman WFI Data Products: What to Expect and How to Access Them**

---

Gisella De Rosa (Mission Scientist at the Science Operation Center)  
Patrick Lowrance (Lead at the Science Support Center)

July 15, 2025



## Roman WFI Data Levels

---

<b>L0</b>	Raw, packetized data	Only available in MAST to authorized accounts
<b>L1</b>	Uncalibrated WFI exposures	3-D WFI non-destructive up-the-ramp samples
<b>L2</b>	Calibrated rate images	Corrected for detector effects, collapsed along time
<b>L3</b>	Resampled co-added images	Corrected for geometric distortion
<b>L4</b>	High-level, extracted information	Associated to L2 or L3 products. Includes source catalogs, light curves, 1-D spectra etc.
<b>L5</b>	Community-contributed products	



## Who Is Responsible for What? SOC vs SSC Products

### Science Operations Center (SOC):

- WFI data cubes and calibrated rate images (L1, L2)
- Resampled co-added images (L3)  
No Galactic Bulge Time Domain
- **Catalogs and high-level products (L4)** associated with SOC-generated L2 and L3 data, includes:
  - Source catalogs
  - Photometry catalogs for variable sources
  - Libraries of empirical PSFs

### Science Support Center (SSC):

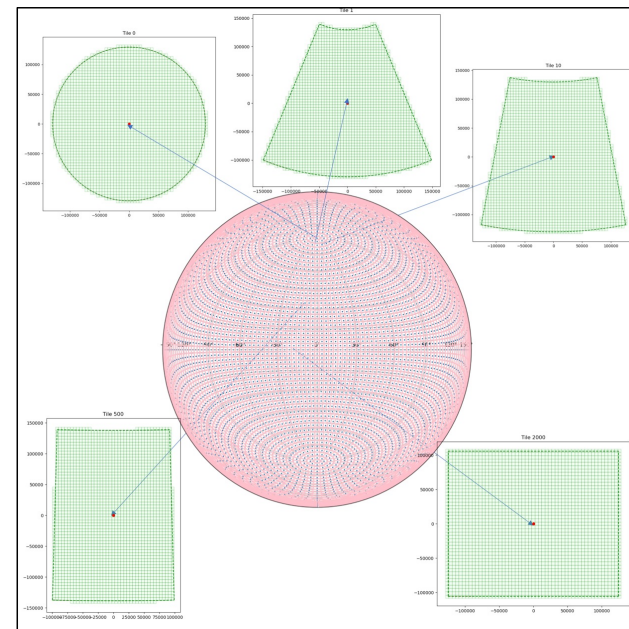
- Resampled, co-added images (L3) and catalogs (L4) associated with **Galactic Bulge Time Domain observations**, includes:
  - Object and Light curve catalogs
  - Variability catalogs
  - Microlensing Object and Event Catalogs
- **Spectroscopic Level 4 products**, includes:
  - 2D and 1D spectra
  - Catalogs containing spectroscopic measurements

All products will be archived in MAST, and user support is integrated



## WFI Data format

- WFI Images and spectra stored in the **Advanced Scientific Data Format (ASDF)**
- Catalogs stored in **Parquet** format
- Roman SOC uses a **tessellation scheme** (HealPix, double pixelization) to partition the sky and store large L3 mosaic
- **L3 products are stored as skycells** - square tiles 4.6 arcmin on a side
- **Skycells are grouped into projection regions (3.2 deg<sup>2</sup>)** - share common WCS, consistent astrometric alignment. Square near the equator, distorted toward the poles.





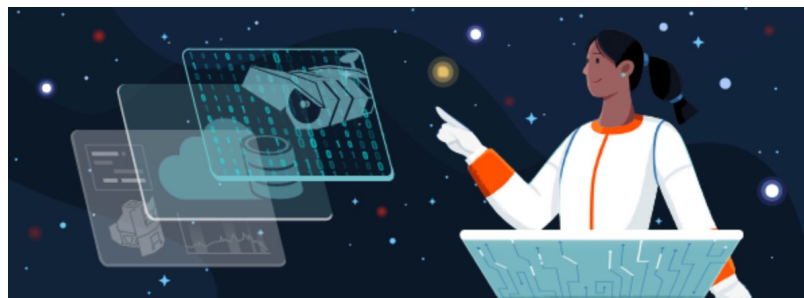
## Prompt Products vs Data Releases

### Prompt Products:

- Processed upon data receipt
- Calibrated using best available reference files and pipeline version at the time - quality may vary
- Not versioned - **overwritten** during subsequent reprocessing

### Data Releases (DR):

- Reprocessed on a **release schedule**, ~6 months after necessary observations are collected
- Calibrated with **uniform pipeline and reference file versions**
- Versioned and archived





## Prompt Products Overview

---

Prompt data products are available in MAST as quickly as they are processed,  
with no access restrictions and no proprietary time

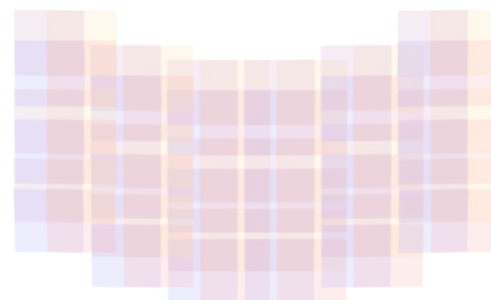
- **Data cubes (L1):** within ~24 hours from receipt
  - Science data and guide window products
  - Reprocessed and overwritten as needed
- **Calibrated rate images (L2):** within ~48 hours from receipt
  - RomanCal, best available calibrations
  - Reprocessed for DRs with uniform pipeline and reference file
  - Only most recent version archived



## Prompt Products Overview

---

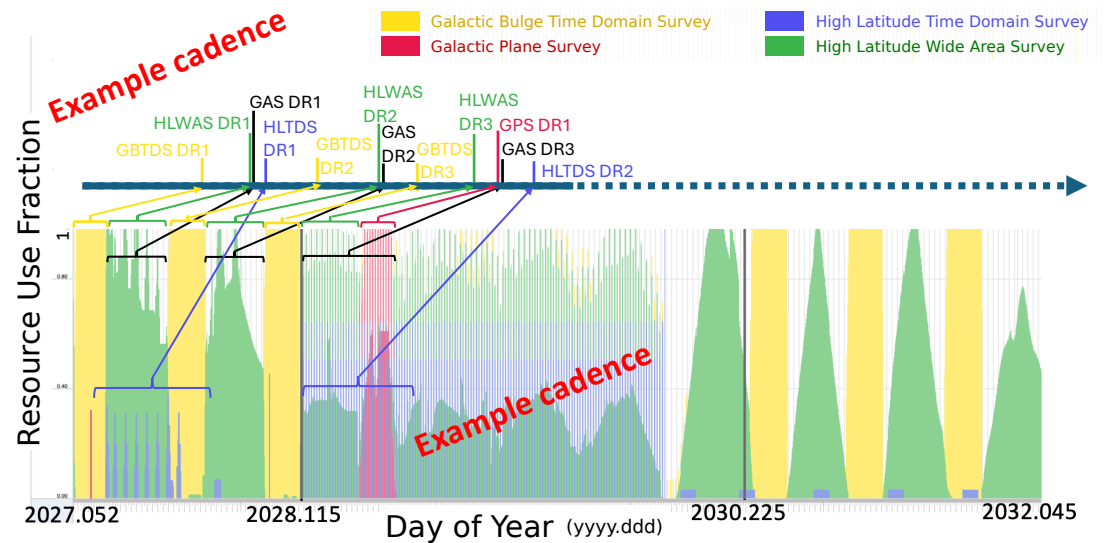
- **Resampled, co-added images (L3):** ~5 days after receipt of last observation
  - Created at **visit level** from prompt L2 products
    - Visit: one or more exposures with consistent instrument settings; may include dithers or repeats
  - Resampled onto Skymap Tessellation
  - Native pixel scale
  - GBTDS L3 products are generated by SSC
- **High level products (L4):** Available ~7 days after receipt of last observation
  - Include single-band source catalogs and segmentation maps from L2 and L3 prompt products
  - SSC generates L4 catalogs for GBTDS and WFI spectroscopy





## Cadence of Data Releases

- DR timing driven by observation schedule - cadences will vary and evolve with mission operations
- Each Community-defined Survey has independent DR
- General Astrophysics DRs may be grouped based on scheduling
- GBTDS/Spectroscopy:
  - SOC: low-level reprocessing
  - SSC: coadded images, spectra, and high-level products



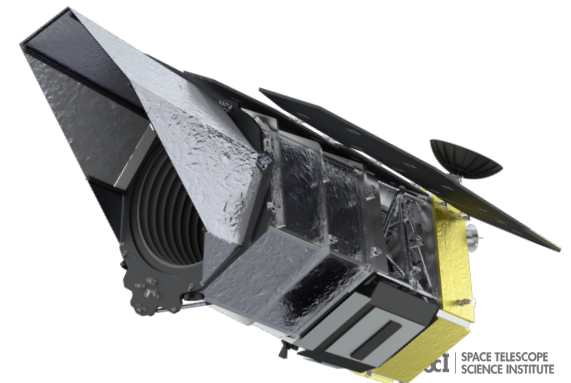


## Overview of SOC Data Release Products

---

DR products are uniformly processed and science-ready  
All DR products are made available simultaneously in the archive

- **Data cubes (L1) & Calibrated rate images (L2):** overwritten if regenerated
- **Resampled, co-added images (L3):** Versioned
  - **DR-SUBSET:** Oversampled image stacks, tied to survey elements
  - **DR-FULL:** Oversampled stack of all the images ever obtained from a program
- **High level products (L4):** Versioned
  - PSF-matched multi-band photometric catalogs
  - Small set of artificial source tests
  - Forced-photometry catalog for time variability
  - Segmentation maps





## How to Access Roman Data

Roman data stored in the Amazon Web Services (AWS) cloud in the US-East 1 data center and available through MAST

Ways to access MAST:

- User interface (UI)
- Python ASTROQUERY package
- Curl scripts – can be generated from MAST UI
- Data Access Application Programming Interface (DAAPI) – GraphQL interface to MAST Archive Catalog



## Roman Research Nexus

---

Roman Observatory will downlink an average of 11 Tb of compressed data per day

	Hubble WFC3	JWST NIRCam	Roman WFI
Number of pixels	1 Megapixel	34 Megapixel	302 Megapixel
Field of view	4.7 arcmin <sup>2</sup>	9.7 arcmin <sup>2</sup>	1035 arcmin <sup>2</sup>
Field of view image data size	0.013 GB (13 MB)	2 GB	8 GB

Need change in paradigm to access and analyze data

**Roman Research Nexus:** low barrier access to data and collaborative resources - compute, storage, and software - to enable and enhance Roman science



## Roman Research Nexus: What is it?

---

### Deployment of JupyterLab environment on Amazon Web Services

Provides users with:

- Cloud data access
- Custom SOC & SSC Software
- Stable software environment
- Accessible and scalable compute resources
- Collaborative design
- Pedagogical framework





## Roman Research Nexus: Collaborative Hub

---

- Shared compute, storage, and egress resources
- Tools to track aggregated resource usage over time
- A variety of computing environments (CPU, RAM)
- Users can belong to multiple teams and access files across all associated teams
- Real-time collaboration features (e.g., simultaneous editing) in beta testing

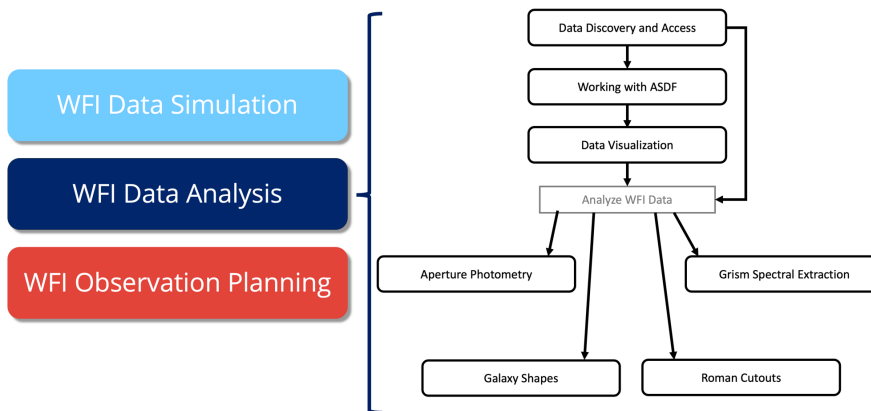
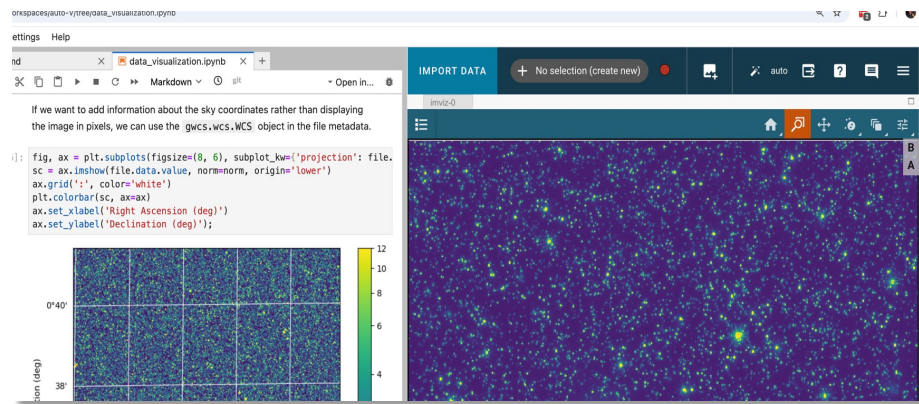




## Science Functionality: Tutorials vs Workflows

### Tutorials

Pedagogical  
Jupyter Notebooks,  
demonstrate how to  
use a tool or perform  
specific task

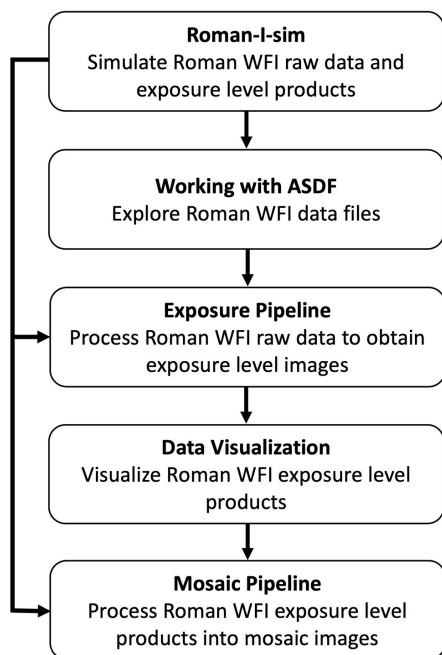


**Science Workflow**  
Combination of  
tutorials and  
documentation that  
guides users through  
a specific science  
use case

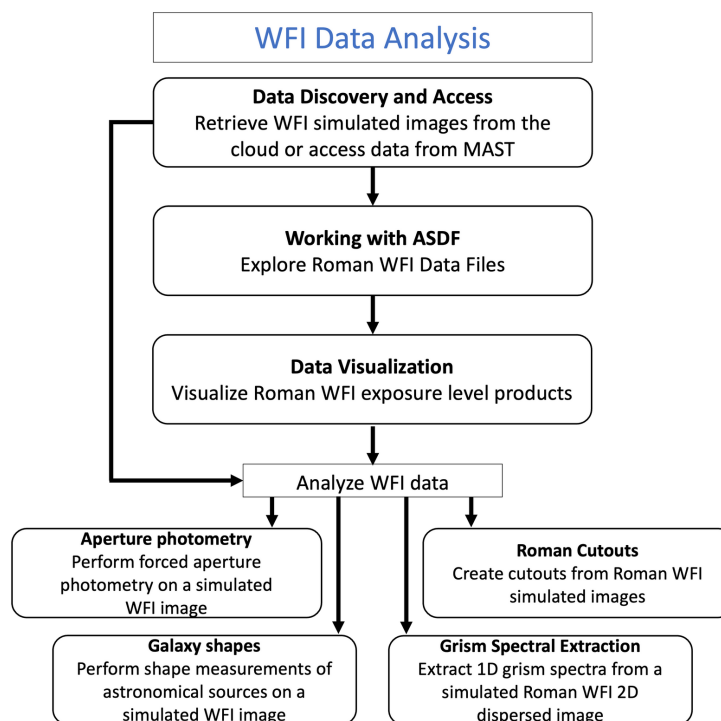


## Science Functionality: Current Offer

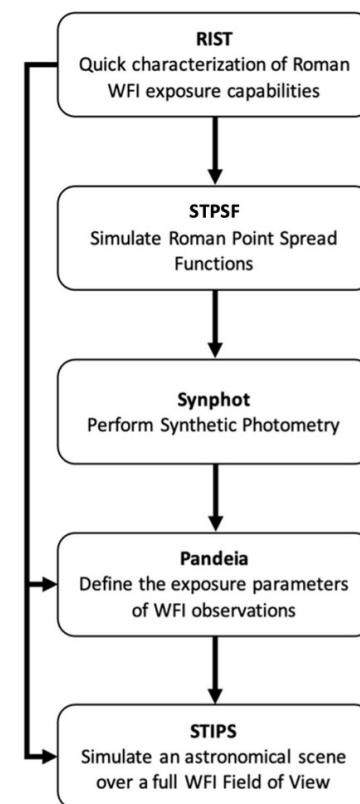
### WFI Data Simulations



### WFI Data Analysis



### WFI Observations Planning





## Development timeline

---

- The RRN is in active development but has already been shared with the community for training and feedback
- **Public release is planned for late Summer to early Fall 2025:**
  - **Individual Accounts**

All users with a MyST account will be able to access the Nexus for basic exploratory work
  - **Team Accounts**

Team accounts automatically assigned to teams with approved Roman programs  
All other teams will be able to request team accounts through simple form
  - **Credit System**

Each account will receive a number of credits to cover AWS-related costs (compute, storage, egress, etc.), with flexibility in usage, and tools to monitor spending



To learn more

---

Roman Documentation System



[roman-docs.stsci.edu](https://roman-docs.stsci.edu)

Roman Help Desk Portal



[romanhelp.stsci.edu](https://romanhelp.stsci.edu)