STScI | SPACE TELESCOPE SCIENCE INSTITUTE

EXPANDING THE FRONTIERS OF SPACE ASTRONOMY

Nancy Grace Roman Space Telescope Data Management and Data Analysis

Henry Ferguson STscI

January 15, 2021

Science Operations and the Data Management System

STScI Roman Science Operation Center responsibilities include:

- Planning & scheduling all observations
- Calibration and support of the Wide Field imaging
- The archive (MAST) for all mission data
  - Most Roman science will be archival due to the survey nature of the mission

NASA Astrophysics *Big Data*:

- Data accumulated per day likely to be >100x *Hubble*
- Both catalogs and pixel-level data sets provide unique science opportunities
- Downloading and processing exceeds resources typically available

Barbara A. MIKULSKI ARCHIVE FOR SPACE TELESCOPES

Data products from multiple mission partners

- Calibrated and mosaiced images, extracted spectra, catalogs, etc.
- Staged in the cloud and co-located with significant computational resources
- Open source, modular imaging pipeline facilitates custom reprocessing

WFI Data Management Environment

- Cloud-based science platform for high-level data processing
- Bring software to the data
- Jupyter Lab environments and notebooks to ease access, sharing and repeatability

AWS

jupyter

STScI | SPACE TELESCOPE SCIENCE INSTITUTE

Roman Data Management

Downlinked science data

STScI
Science Operations Center
Data Management System

WFI Exposure-level processing

WFI High-level processing

Archive

Data Analysis

Low-level processed data

IPAC
Science Support Center

WFI high-level processing
Microlensing & spectroscopy

CGI data analysis environment

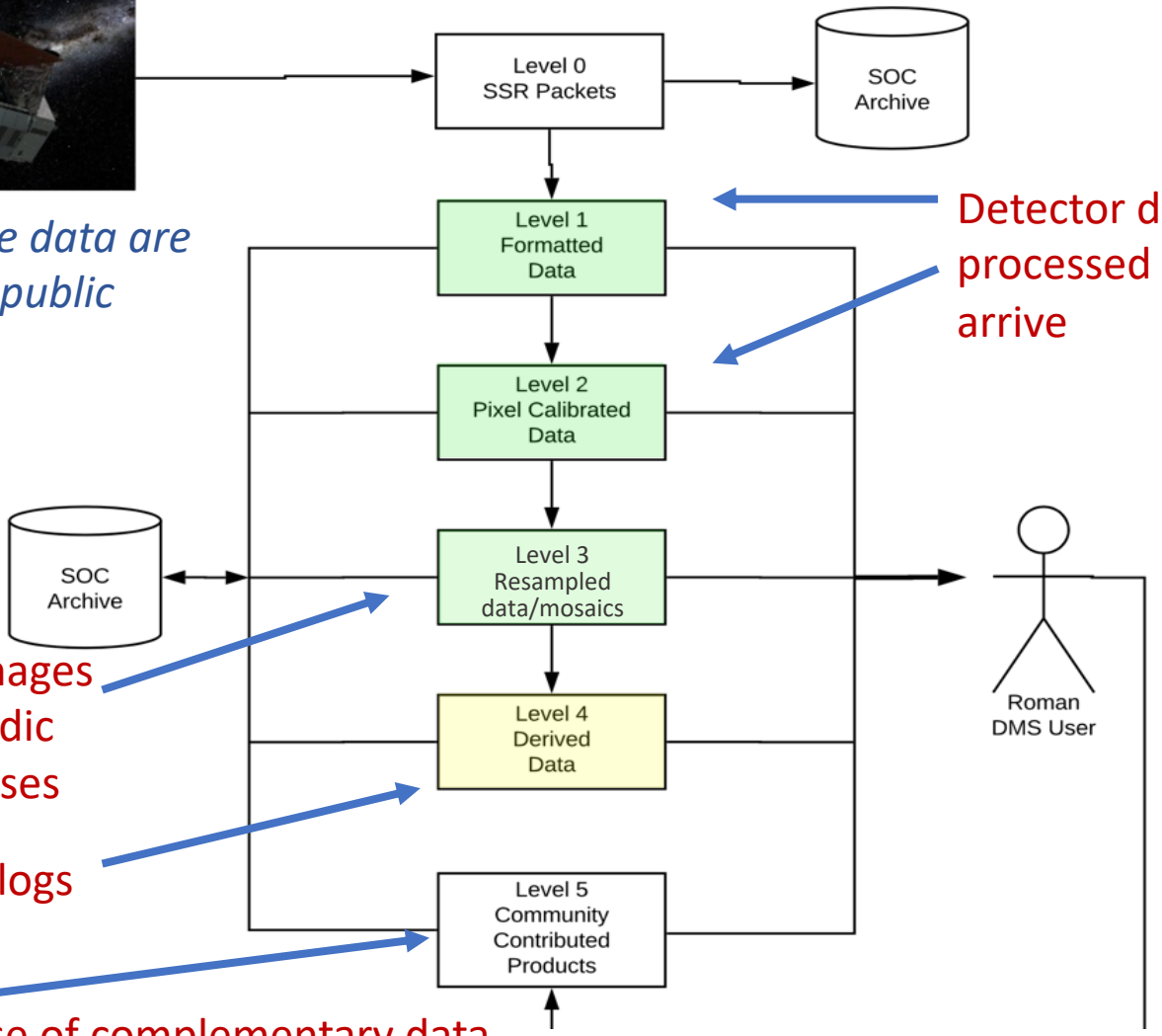
Science ready data

Data and software

Science community

Data product levels

All Roman science data are immediately public



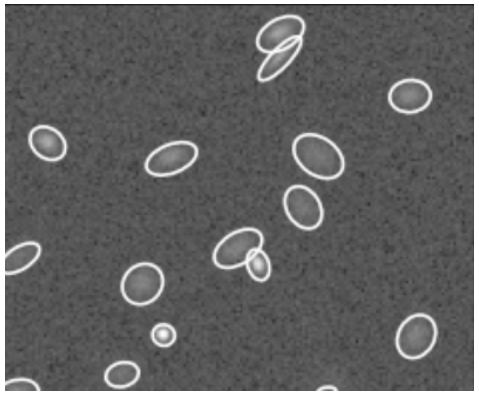
Available as associated images are processed; with periodic consolidated survey releases

Catalogs

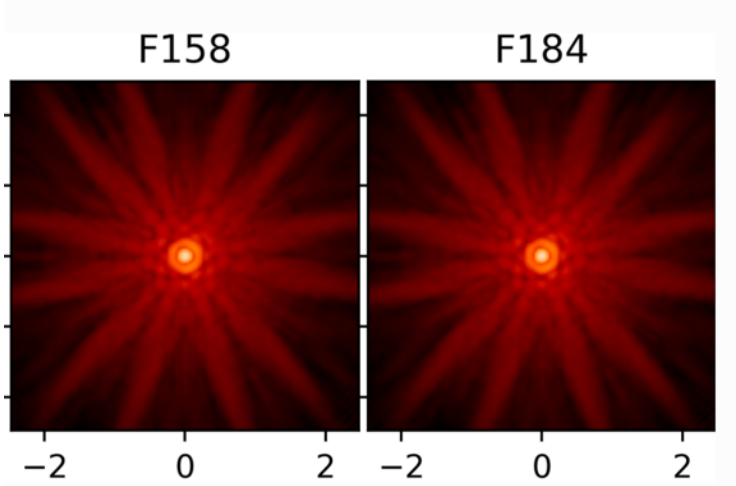
E.g. catalogs that make use of complementary data from other facilities & survey-level calibrations

WFI Imaging
High-Level Pipeline Components

Catalogs



Point spread functions

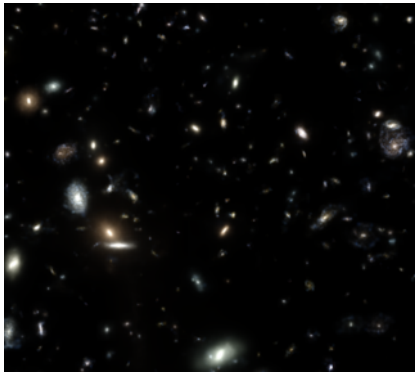


High-level science products:

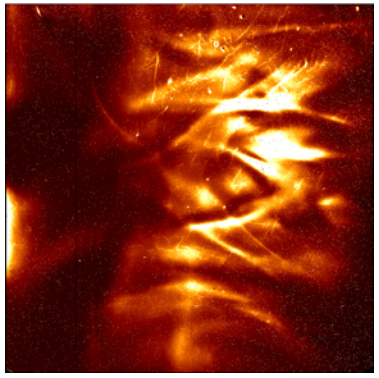
- Catalogs of static & variable sources
- Spectroscopic extractions & redshifts (SSC)
- Photometric redshifts (based only on Roman data)

Simulations

Astronomical Sources



Instrument signatures



→ Completeness/Systematics

Community input

- Moving from concept to design
  - Helpful to have use cases and “user stories” to help guide development
- Types of questions
  - What measurements would you like to see in the catalog(s)?
  - Do you anticipate needing to reprocess image data? Why?
  - Do you anticipate needing to download large amounts of data, instead of working in the cloud? Why?
  - How can we best configure the science platform (e.g. with what hardware or software) to meet your needs?

Email ferguson@stsci.edu