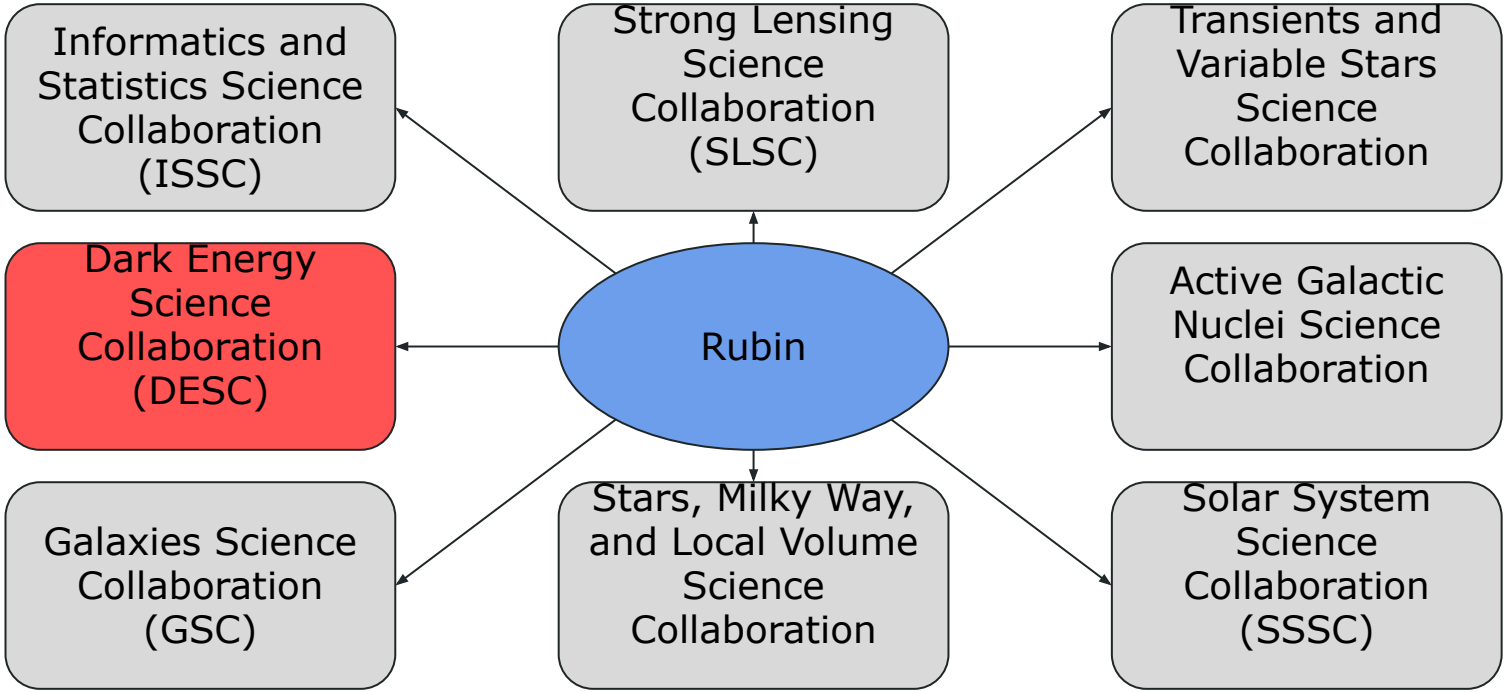
Data Access and Validation

October 2024 Sprint week tutorial

LSST DESC
Dark Energy Science Collaboration

Intro

Rubin and the science collaborations

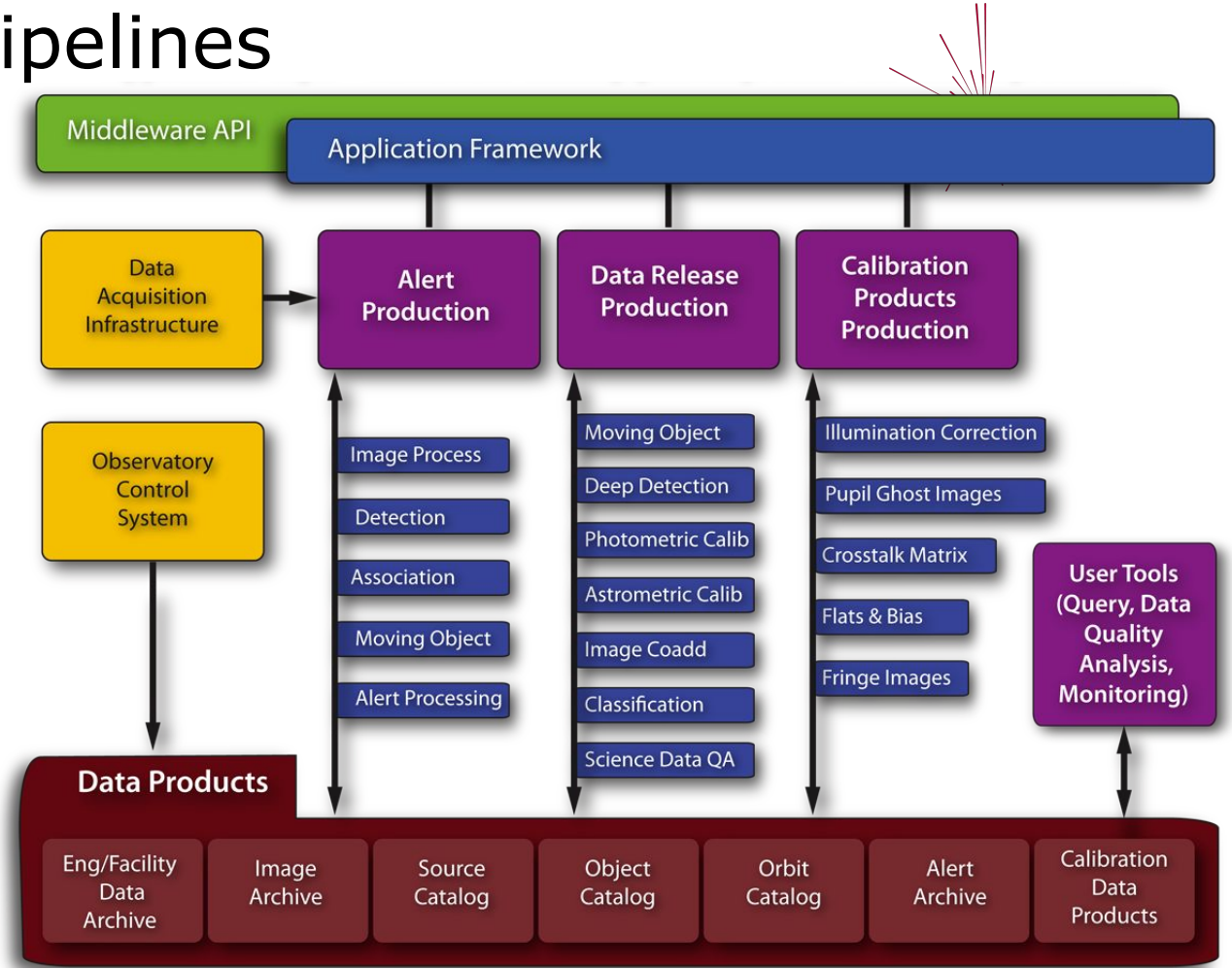


You can find links to individual collaborations at
https://lsstdiscoveryalliance.org/lsst-science-collaborations/

LSST Science Pipelines

https://pipelines.lsst.io/

- Huge amount of pipeline infrastructure to go from observatory images to object catalogs and more
- You'll hear this referred to as the DMStack
- Installed on the

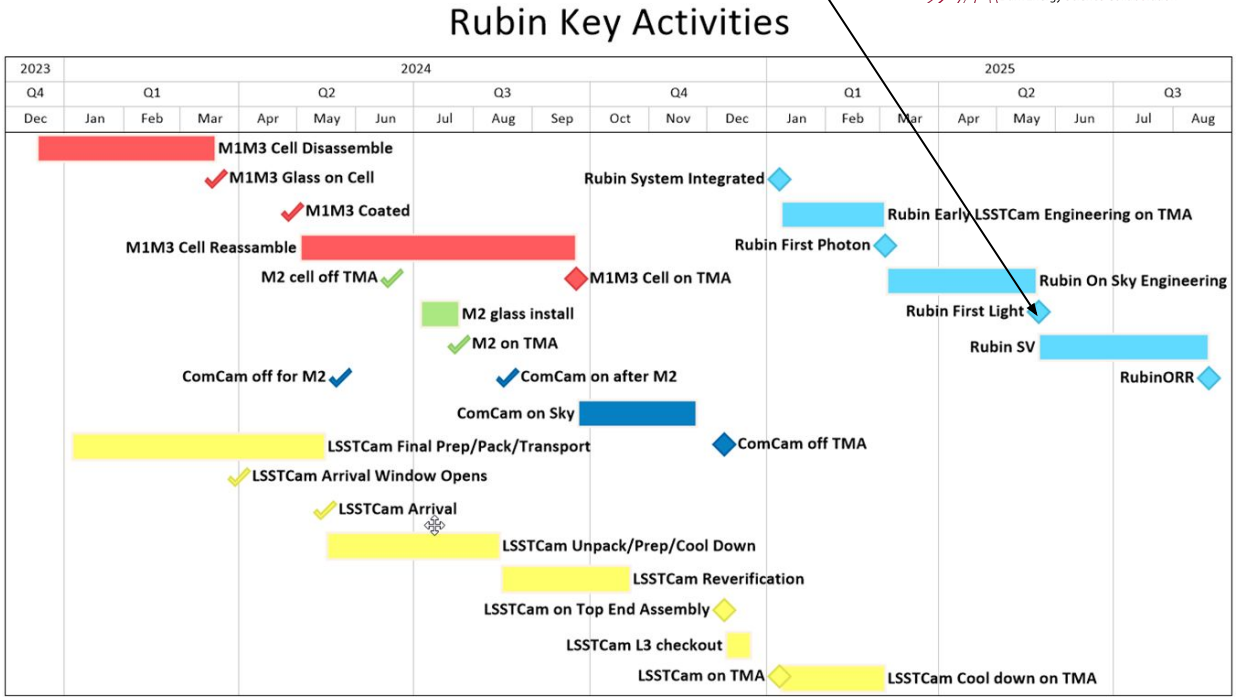


https://www.lsst.org/about/dm/pip

Data timeline

First light mid-2025

- First light June 2025
  - DP1 (first data preview release) 2-3 months later, ~Aug-Sep 2025
  - Survey start 4-7 months later, ~Oct 2025 -Jan 2026
  - DP2 (second data preview) 9-12 months later, ~April-June 2026
- First data release (DR1) anticipated 12-14 months after survey start
  - Roughly late 2026- early 2027



From, https://www.lsst.org/about/project-status, updated Au 2024

(numbers on figure slightly out of date, first light has since b

What will be in the previews?

- We may end up with more than this
- Simulated data, precursor data available in the meantime (e.g. HSC, skysim/CosmoDC2, DC2, DP0.2, roman-rubin)
- For this tutorial we'll be working with DP0.2 data preview. Uses DC2 images processed by the project. Find links and

Rubin Early Science Data Release Scenario

Data Product	Jun 2021 DP0.1 DC2 Simulated Sky Survey	Jun 2022 DP0.2 Reprocessed DC2 Survey	Jun 2023 DP0.3 Solar System PPDB Simulation	Oct 2024 – Jul 2025 DP1 ComCam or early LSSTCam Data	Nov 2025 – May 2026 DP2 LSSTCam Science Validation Data	May 2026 – Jan 2027 DR1 LSST First 6 Months Data	May 2027 – Jan 2028 DR2 LSST Year 1 Data	May 2028 – Nov 2028 DR3 LSST Year 2 Data
Raw Images	●	●	–	●	●	●	●	●
DRP Processed Visit Images and Visit Catalogs	●	●	–	●	●	●	●	●
DRP Coadded Images	●	●	–	–	●	●	●	●
Object and ForcedSource Catalogs	●	●	–	–	●	●	●	●
DRP Difference Images and DIASources	–	●	–	–	●	●	●	●
DRP ForcedSource Catalogs including DIA output	–	●	–	–	●	●	●	●
PP Processed Visit Images	–	–	–	–	–	●	●	●
PP Difference Images	–	–	–	–	–	●	●	●
PP Catalogs	–	–	–	–	●	●	●	●
PP SSP Catalogs	–	–	●	–	●	●	●	●
DRP SSP Catalogs	–	–	–	–	–	●	●	●

TABLE 1: Summary of data products expected in each data preview and early survey data release.

https://rtn-011.lsst.io/

Data on NERSC

Current plan is approximately

DESC-generated products +

- coadd object catalogs as flat files
- metadetection shear catalogs as flat files
- cell-based coadds for the main survey
- cell-based coadds for the deep-drilling fields
- survey masks
- survey property maps
- catalogs of stars used for the PSF model fits w/ moments measurements for those stars and the models of them
- PSF models for the images (visits)
- astrometric solutions for the images
- smaller amounts of final visit images for TD analysis
- small amounts of raw data w/ butler for reprocessing tests
- visit metadata / AuxTel information

Data required for key analyses will be transferred to NERSC so we can use DESC compute resources.

For everything else you would need to access data on the RSP

If you see anything missing please contact Matt Becker or Nacho Sevilla ASAP

Contents

- The Rubin Science Platform
- Accessing DP0.2 at NERSC
- Validating the data
- Data inspection/ validation challenges

10 minute DIY task/break time between each section

Working on the Rubin Science Platform

LSST DESC
Dark Energy Science Collaboration

Intro to the RSP

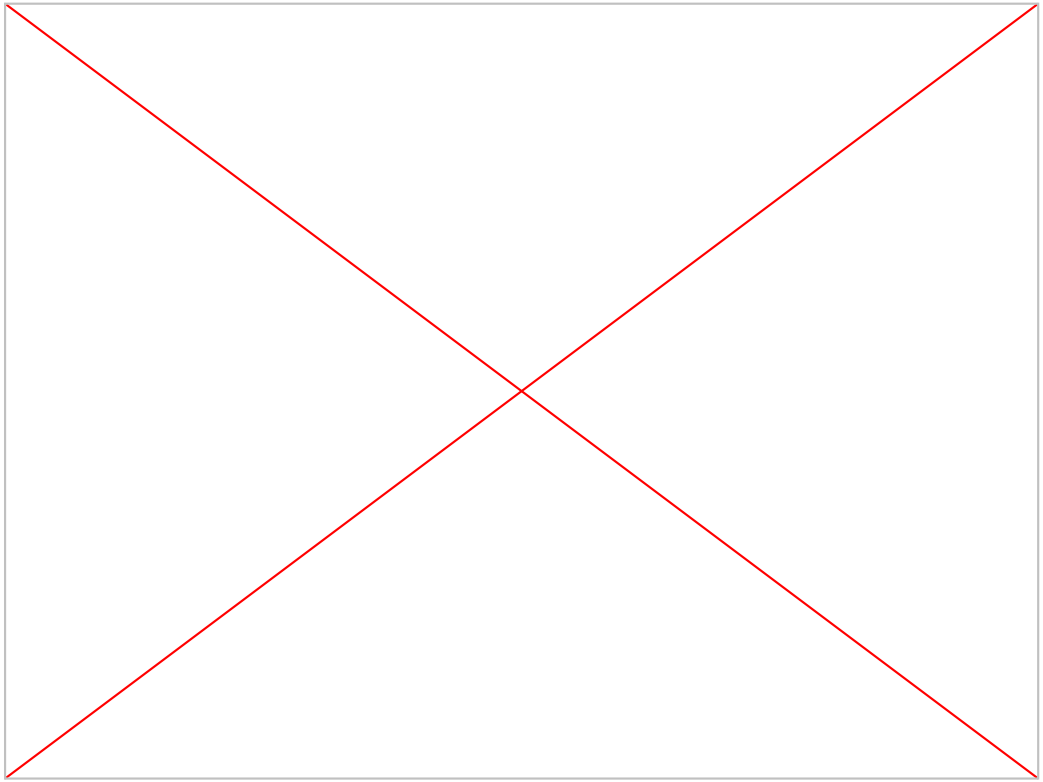
Get a delegate account

- https://dp0.lsst.io/delegate-resources/index.html
  Quote: There is no application process. The only prerequisite is to be a Rubin data rights holder (see the Rubin Data Policy). The number of delegates is currently limited to 900. Space is available and all petitions are being accepted. If that changes this page will be updated.

Log in to the RSP

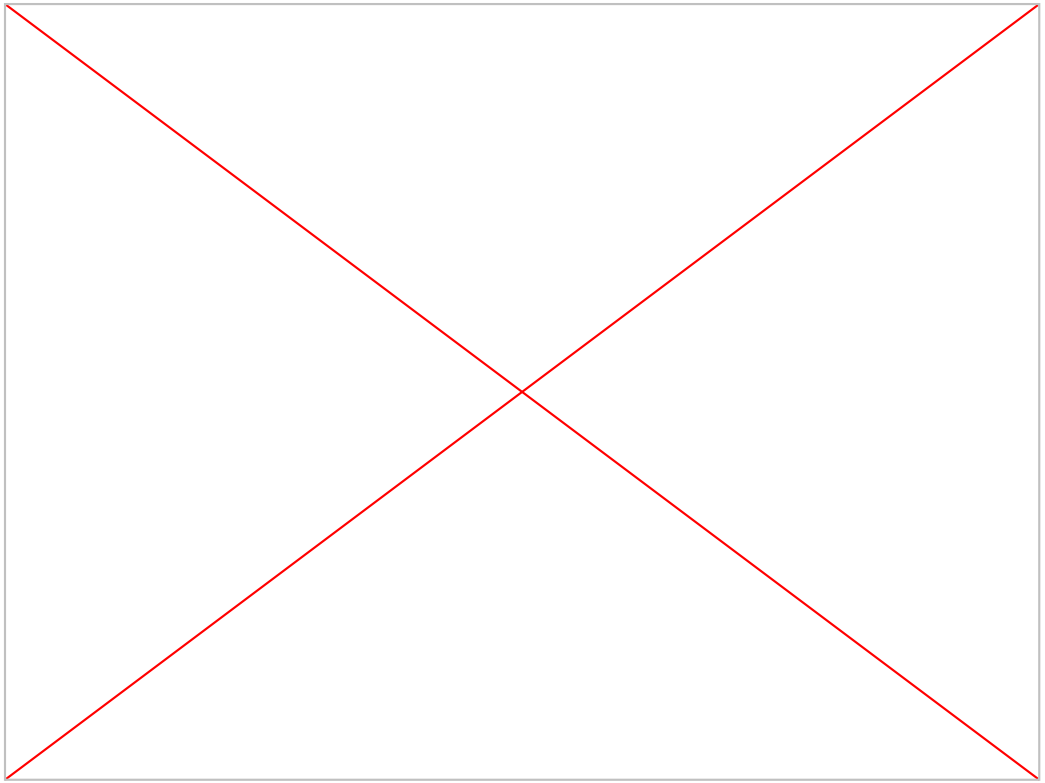
- https://data.lsst.cloud/
  (Live demo. Links to follow along here: https://dp0-2.lsst.io/tutorials-examples/portal-beginner.html https://dp0-2.lsst.io/_static/nb_html/DP02_01_Introduction_to_DP02.html)

Intro to the RSP - portal demonstration



https://dp0-2.lsst.io/tutorials-examples/portal-beginner.html

Intro to the RSP - notebook demonstration



- https://dp0-2.lsst.io/_static/nb_html/DP02_01_Introduction_to_DP02.html

Intro to the RSP

Tutorials home:
https://dp0-2.lsst.io/tutorials-examples/index.html

Notebook tutorials:

https://github.com/rubin-dp0/tutorial-notebooks

A couple of particularly DESC-relevant notebooks on PSFs:

- https://github.com/rubin-dp0/tutorial-notebooks/blob/main/DP02_12a_PSF_Data_Products.ipynb
- https://github.com/rubin-dp0/tutorial-notebooks/blob/main/DP02_12b_PSF_Science_Demo.ipynb

Intro to the RSP

Issues? Try the community forum

- https://community.lsst.org/

More info about dp02?

- https://lsstdesc.org/srv-dp02/tutorials/Rubin/DP02_intro.html

Quick RSP exploration! - 10 minute pause

Options:

- Apply for DP0 delegate position
- Explore the RSP tutorials
- Ask questions
- Have a coffee/bathroom break!

Moving to jupyter!

- Log in to jupyter notebooks on NERSC

  > https://jupyter.nersc.gov/

- Git clone this repo (You can select File→New→Terminal to open up a terminal window, and in your home directory type)

  > git clone https://github.com/patricialarsen/desc_data_tutorial_2024.git

- Set up your kernels (in the same terminal window type the following)

  > source /global/common/software/lsst/common/miniconda/kernels/setup.sh

- Open tutorial in a notebook with the desc-python kernel