Vera C. Rubin Observatory
Rubin Observatory Operations

Data Preview 1: Definition and planning

William O'Mullane

RTN-085

Latest Revision: 2025-02-06

D R A F T

Abstract

This document is to centralise information and planing for DP1. It will point to other documents for details.

Change Record

Version	Date	Description	Owner name
1	YYYY-MM-DD	Unreleased.	William O'Mullane

Document source location: `https://github.com/lsst/rtn-085`

Contents

Data Preview 1: Definition and planning

1 Introduction

We pull together information from Jira and refer to other documents as needed to describe DP1 in more detail.

DP1 should be released to the community after the First Look image is released to the press. The general order with some time (days potentially weeks) between steps is:

- Public First Look - JPG and large images full focal plane or more

- DP1 - ComCam data reprocessed and in good form

- Science Images in the FL images

1.1 DP1 Scope

The goal od DP1 is to give data to the science community - not any new features in RSP etc.

1.1.1 Data

DP1 comprises data from ComCam on sky. RTN-011 gives an outline of DP1 however at time of writing it still states DP1 could be either camera.

DP1 could contain data with poor PSF as ComCam may be driven by work on AOS verification. As such:

- There will be PVIs and catalogs if the data are good enough.

- There is no promise of coadds but if there is good data this is a stretch goal.

- There will be no difference imaging in DP1.

1.1.1.1 First Look Images   First Look images from LSSTCam will not be part of DP! but should follow sometime after it. The FL will only be what as done for producing the images of which there should be a small number. It should have PVIs PSF matched coadds and may have catalogs if it is feasible. DM currently have no visibility on what kind of imaging will be taken so can not commit to anything for FL.

1.1.2 Science Platform

Science Platform functionality for DP1 is focusing on good user experience and system scalability rather than full feature capability. For reputational and morale reasons, it is critical that it is adequately communicated to the community that this is not just a data preview but a science platform preview.

Due to the expected low interest in the simulated data based DP0 releases, and due to the decision to open DP1 to all data rights holders, DP1 could be a two orders of magnitude scale-up event over current usage and the first real test of the hybrid model. Preparing for such a jump in usage is challenging and takes priority even over releasing features that we would normally consider part on a minimal feature set.

Nevertheless, some capabilities that were not offered in DP0 nevertheless have to be offered in DP1 as not doing so falls well beyond the minimum standard the community will expect. These are:

- Temporary table upload for Qserv

- User query history

- User query management (requires the new Qserv front end)

- RA/DEC cutouts across patch/tract boundaries

- HIPS for full focal plane images

- Additional data-link annotations

- Butler Client-server

- Visit table in Qserv

- Portal-accessible user WebDAV service

- User quotas

As is evident, some of these require features in development from Qserv and Butler. Additionally documentation activities, helpdesk arrangements, and review of tutorial notebooks for performance will be co-ordinated with the Community Science Team.

It is unlikely that any other capabilities will be ready and hardened for large spiked usage. This means that a number of features of the final system such as bulk cutouts, interactive full focal plane scale viewer, SIA v2 service, persistent user table uploads, consolidated Database, Qserv intersects and a PSF service are NOT expected to participate in DP1 in the current schedule.

2 Planning

We have some milestones and epics leading up to DP1.

We are missing some like Qserv - loading ..

Dates are wrong ..

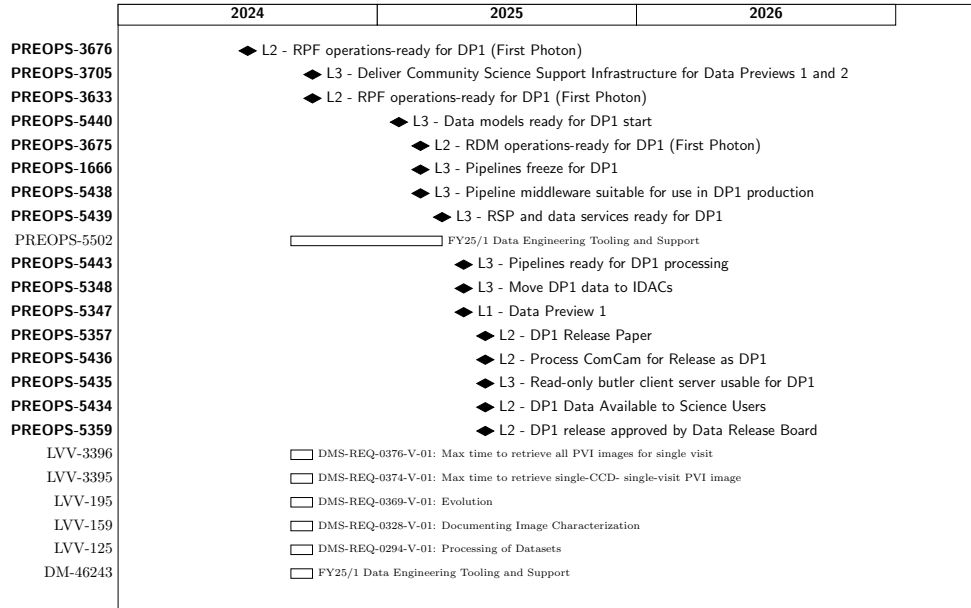Label milestones and epics DP1 to have them appear

FIGURE 1: Data Preview 1 - plan

Open milestones are listed in Table 1.

Table 1: Milestones for Rubin Observatory Data Management and System Performance

Milestone	Jira ID	Rubin ID	Due Date	Level	Status	RubinTeam
L3 - Data models ready for DP1 start	PREOPS-5440	not set	2025-01-31	3	To Do	Data Engineering
L2 - RDM operations-ready for DP1 (First Photon)	PREOPS-3675	not set	2025-02-19	2	To Do	RDM Management
L3 - Pipelines freeze for DP1	PREOPS-1666	L3-AP-0030	2025-02-26	3	To Do	Algorithms & Pipelines
L3 - Pipeline middleware suitable for use in DP1 production	PREOPS-5438	not set	2025-02-28	3	To Do	Pipeline Middleware
L3 - RSP and data services ready for DP1	PREOPS-5439	not set	2025-03-19	3	To Do	Service Quality and Reliability Engineering
L3 - Pipelines ready for DP1 processing	PREOPS-5443	not set	2025-04-15	3	To Do	Algorithms and Pipelines
L3 - Move DP1 data to IDACs	PREOPS-5348	not set	2025-04-21	3	To Do	US Data Facility
L1 - Data Preview 1	PREOPS-5347	not set	2025-04-30	1	Invalid	None
L2 - DP1 Release Paper	PREOPS-5357	not set	2025-05-01	2	To Do	System Performance Management
L2 - Process ComCam for Release as DP1	PREOPS-5436	FY25-2.3.1-M02	2025-05-15	2	To Do	RDM Management
L3 - Read-only butler client server usable for DP1	PREOPS-5435	not set	2025-05-15	3	To Do	Pipeline Middleware
L2 - DP1 Data Available to Science Users	PREOPS-5434	FY25-2.3.1-M01	2025-05-15	2	In Progress	RDM Management
L2 - DP1 release approved by Data Release Board	PREOPS-5359	not set	2025-05-15	2	To Do	System Performance Management

2.1   Component plans

2.1.1   Middleware

- IVOA-compatible SIA image service

- Qserv query temporary uploads

- User query history capabilities

- Context-aware documentation

- Improved notebook-portal integration

There will be no major computational expansion (bulk cutouts, parallel computing, batch processing) is available at this time.There will be quota and API limit handling which along with the roadmap will soon appear on `rsp.lsst.io`.

The milestone for RSP readiness (PREOPS-5439) is some months before DP1 to allow tutorial testing etc.

2.2.0.1 Qserv Data loading.

2.3 Rucio

What do we need for data to be sent to DACs?

3 Processing

As always the goal is to generate a fully self-consistent data release for the scientists to publish papers on.

Mini runs will be done prior to the final production run prior to the pipelines freeze PREOPS-1666.

Who is piloting the DP1 campaign ? PFC ?

3.0.1   Policy committee

There are certain decisions which will need to be made are best handled in a smaller forum than the full leadership team. As fopr previous DPs this may include:

- Campaign polices

- Version of pipelines to use and patches which are needed

- Version of QA tooling which needs to run (and where/how to run it)

- Other operational considerations

Such decisions will be endorsed by DMLT but advised by a smaller committee more connected to the issues. The members will be the following (or their delegated representative):

- Hsin-Fang Chiang

- Tim Jenness

- Yusra AlSayyad

- Colin Slater

3.0.2   Science pipelines release

We have milestone PREOPS-1666 for the pipeline freeze. We can estimate already that this will be V25 of the pipelines and will be driven by the DP1 needs - not on a fixed date.

3.0.3   High level workflow of workflows

Anything to say about this here ?

4   Validation and Documentation

System performance will lead validation of the data for DP1.

Are we having dp1.lsst.io for documentation ? (Assume so) .. Milestones

4.1 Data release board

The data release board should be in place to ok the DP1 at selection and publication citertn-052.

A References

[RTN-011], Guy, L.P., Bechtol, K., Bellm, E., et al., 2024, Rubin Observatory Plans for an Early
Science Program, URL `https://rtn-011.lsst.io/`,
Vera C. Rubin Observatory Technical Note RTN-011

B Acronyms

Acronym	Description
AOS	Active Optics System
AP	Alert Production
API	Application Programming Interface
CCD	Charge-Coupled Device
DEC	Declination
DM	Data Management
DMLT	DM Leadership Team
DMS	Data Management Subsystem
DMS-REQ	Data Management System Requirements prefix
DP	Data Production
DP0	Data Preview 0
DP1	Data Preview 1
FL	First Look
FY24	Financial Year 24

FY25	Financial Year 25
HIPS	Hierarchical Progressive Survey (IVOA standard)
IDAC	Independent Data Access Center
IVOA	International Virtual-Observatory Alliance
JPG	Joint Photographic Experts Group
L1	Lens 1
L2	Lens 2
L3	Lens 3
LVV	LSST Verification and Validation
OPS	Operations
PSF	Point Spread Function
PVI	Processed Visit Image
QA	Quality Assurance
RA	Risk Assessment
RDM	Rubin Data Management
RPF	Rubin system PerFormance
RSP	Rubin Science Platform
RTN	Rubin Technical Note
SIA	Simple Image Access (IVOA standard)
US	United States