

Data Preview 2: Definition and planning
Yusra ALSayyad, and William O'Mullane
2025-12-11

1 Introduction

Between April and September 2025, LSSTCam acquired a sequence of science-grade commissioning images, including small-field observations and the Science Validation surveys. Data Preview 2 (DP2) will consist of a processed set of these observations and will provide early access to science-grade LSSTCam data products to support preparation for LSST science in advance of Data Release 1. This document describes the scope of DP2 and the plan for producing and serving its data products. The intended audience is Rubin staff to use as reference on the DP2 dataset and delivery plan.

DP2 will include images and catalogs from LSSTCam data only (i.e. no AuxTel or LSSTComCam data will be included). Prompt data product releases, alerts, and related services expected in 2026 are out of scope for this document and will be described elsewhere.

Expected release of DP2 to the data rights holders is mid-2026.

2 Dataset Definition

Data Preview 2 (DP2) will include LSSTCam observations acquired between 20 April 2025 and 21 September 2025. It will also include any observations taken between 20 October 2025 and 1 January 2026 that overlap the Science Validation regions. Within this temporal range, all visits labeled as science observations will be included. Engineering data, out-of-focus sequences acquired for active optics commissioning, and stray-light scans will not be included. Figure 1 shows the resulting sky footprint.

The data products produced for DP2 will follow the categories defined in Table 6 of RTN-011. These include calibrated and coadded image products, associated catalogs, and the corresponding calibration products. A complete enumeration of DP2 data products will be provided later in this document.

No quality criteria will be applied to determine whether a visit is included in DP2. However, quality criteria will instead determine whether a visit contributes to the coadds. Visits will be included in the coadds only if the visit-level median PSF FWHM is better than 1.7 arcseconds, if each detector has a FWHM better than 1.8 arcseconds and a transparency of at least 0.75. Out-of-focus data will naturally be excluded from the coadds because the pipelines fail to model the PSFs of donut images.

2.1 Visit-level and Detector-level exclusions

The eight heavily vignettted corner detectors (0, 20, 27, 65, 123, 161, 168, 188) will be excluded. Detectors 172, 24, 120, 121 will be excluded for data taken before 3 July 2025 because the data to calibrate them does not exist.

Only visits labeled “science” will be included. Exposures labeled “acq”, “engtest”, and “cwfs” will not be included. These images were taken to test hardware and are often intentionally out of focus. Table 1 summarizes the additional visit-level exclusions:

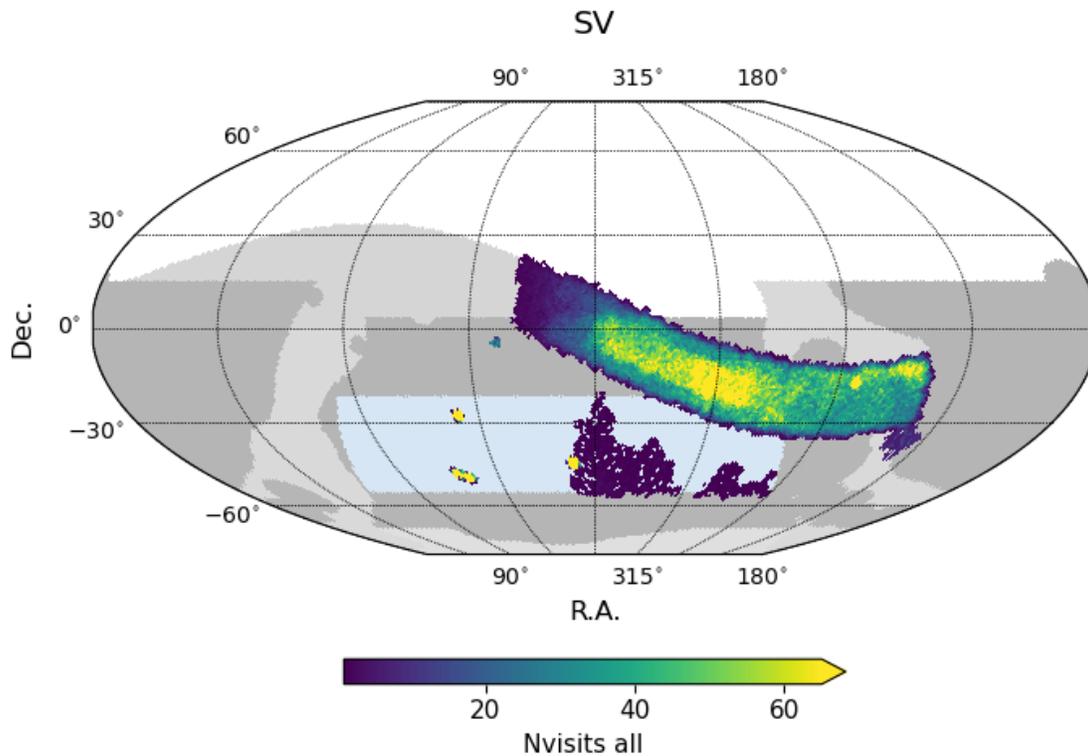


FIGURE 1: Sky footprint of the DP2 dataset. Coadds will include the ecliptic stripe and deep drilling fields.

TABLE 1: Summary of visit-level exclusions.

Excluded visits	N visits	details
Visits taken before day_obs 20250424	556	
Visits taken without a filter (u, g, r, i, z, y)	1	
Stray light scans	305	BLOCK-T516, BLOCK-T519, BLOCK-T540
Pinhole images	1	BLOCK-T491
AOS Empirical Sensitivity Performance	2	BLOCK-T545
M1M3 Thermal Step Test	581	BLOCK-T609
all_sky_aos_test	772	Initial AOS survey test, mostly in <i>y</i> -band, taken in June.
vignette in (FULLY, PARTIALLY, UNKNOWN)	67	
Bad visits in excluded_visits list	391	Visually flagged errors, typically out-of-focus images or exposures affected by mount or hexapod motion.

Filtering on the above criteria results in a total of 21,067 visits, split into the fields and bands shown in Table 2. Not all of these meet the quality criteria for inclusion in the deep coadds: $\text{eff_time_zero_point_scale_median} > 0.75$ and per visit PSF FWHM < 1.7 arcseconds. The 15,610 visits that do meet these criteria are broken out in Table 3.

2.2 Data taken after Oct 2025

Data taken after Oct 2025 may be included pending the generation of suitable calibration frames for that epoch before processing starts. If suitable calibration frames are made before processing, intent is to include data that covers the region in Figure 1 in DP2. Inclusion of single-frame data products (Visit Images and Sources) for all other “science” visits (BLOCK-407, BLOCK-408, BLOCK-416) which are unsuitable for coadds may also be considered.

3 Data Production

DP2 processing will be based on Version 30 of the LSST Science Pipelines. A minor release will be created at the start of each processing stage. Backports required for a given stage must be submitted two weeks prior to the start of that stage to allow time for validation. A two-week pilot run for each stage will be executed at the IN2P3 Data Facility.

TABLE 2: Total visits by target and band

Target	u	g	r	i	z	y
lowdust, bulgy or dusty_plane	709	1313	2550	3130	2768	2147
COSMOS	100	82	166	139	110	66
DDF ECDFS	0	36	39	41	43	0
DDF EDFS_a	0	19	21	28	18	0
DDF EDFS_b	0	23	20	29	17	0
DDF ELAISS1	39	101	102	163	112	16
ELAIS_S1	11	30	30	30	35	30
M49	237	280	378	214	0	0
New_Horizons	36	48	70	108	74	23
Prawn	196	164	142	93	30	0
Rubin_SV_212_-7	0	139	236	123	0	0
Rubin_SV_216_-17	0	28	30	63	0	0
Rubin_SV_225_-40	312	568	441	381	240	104
Rubin_SV_280_-48	30	30	29	30	29	0
Rubin_SV_300_-41	0	8	0	0	30	0
Rubin_SV_320_-15	0	29	17	105	52	70
ToO_3l_i0	0	16	14	14	7	0
ToO_BBH_i0	18	0	18	26	0	0
ToO_SSO_i0	0	16	16	16	12	16
Trifid-Lagoon	235	196	122	115	0	0
nes	6	42	175	168	209	39
<i>Insufficient visits to coadd</i>						
Abell_2764	0	7	0	0	0	0
DDF XMM_LSS	30	0	0	0	0	0
DESI_SV3_R1	0	0	0	4	0	0

DP2 production will follow the four-stage structure used for data release processing. Stage 1 consists of single-visit processing. Stage 2 consists of recalibration and will include FGCM for photometric calibration, GBDES for astrometry, Gaussian-process modeling of atmospheric turbulence, proper motion and parallax fitting, and PSF re-estimation using a standard star set. Stage 3 consists of coaddition and measurement on coadds, including the generation of object tables. Stage 4 consists of difference-image analysis, DIA Source and DIA Object generation, Solar System object processing, and the generation of light curves via the forced-source tables.

DP2 production will be divided between the US Data Facility (USDF) at SLAC and the French data facility at IN2P3. Workload distribution is expected to be approximately equal between USDF and IN2P3. Figure 2 illustrates the planned distribution of processing activities across the data facilities. Each stage of processing will include continuous monitoring of data quality

TABLE 3: Deep coadd quality visit counts by target and band

Target	u	g	r	i	z	y
lowdust, bulgy or dusty_plane	531	913	1662	2182	2108	1818
COSMOS	56	71	124	98	83	51
DDF ECDFS	0	10	27	30	25	0
DDF EDFs_a	0	9	10	22	10	0
DDF EDFs_b	0	7	8	24	10	0
DDF ELAISS1	13	53	51	123	64	0
ELAIS_S1	0	0	0	29	34	29
M49	41	235	361	209	0	0
New_Horizons	21	47	70	77	52	23
Prawn	133	148	132	93	27	0
Rubin_SV_212_-7	0	112	217	123	0	0
Rubin_SV_216_-17	0	28	30	63	0	0
Rubin_SV_225_-40	159	324	265	328	207	100
Rubin_SV_280_-48	1	27	27	30	27	0
Rubin_SV_300_-41	0	1	0	0	20	0
Rubin_SV_320_-15	0	2	3	101	48	70
ToO_3I_i0	0	8	14	13	7	0
ToO_BBH_i0	9	0	18	16	0	0
ToO_SSO_i0	0	16	16	16	12	16
Trifid-Lagoon	232	191	122	115	0	0
nes	1	26	144	128	180	33
<i>Insufficient visits to coadd</i>						
Abell_2764	0	7	0	0	0	0
DESI_SV3_R1	0	0	0	3	0	0

and the relevant verification and validation metrics. A full verification and validation review is required at the end of each stage before processing may advance to the next stage.

The Data Production division of Data Management is responsible for generating the data products that comprise DP2, after which the dataset will be transferred to the Data Services division for user-facing delivery. Data Production includes the Science Pipelines, Campaign Management, and Verification and Validation teams.

The Science Pipelines team is responsible for the development and configuration of the software used in DP2 processing. The Campaign Management team is responsible for planning and executing the production, managing workflow execution across facilities, and verifying

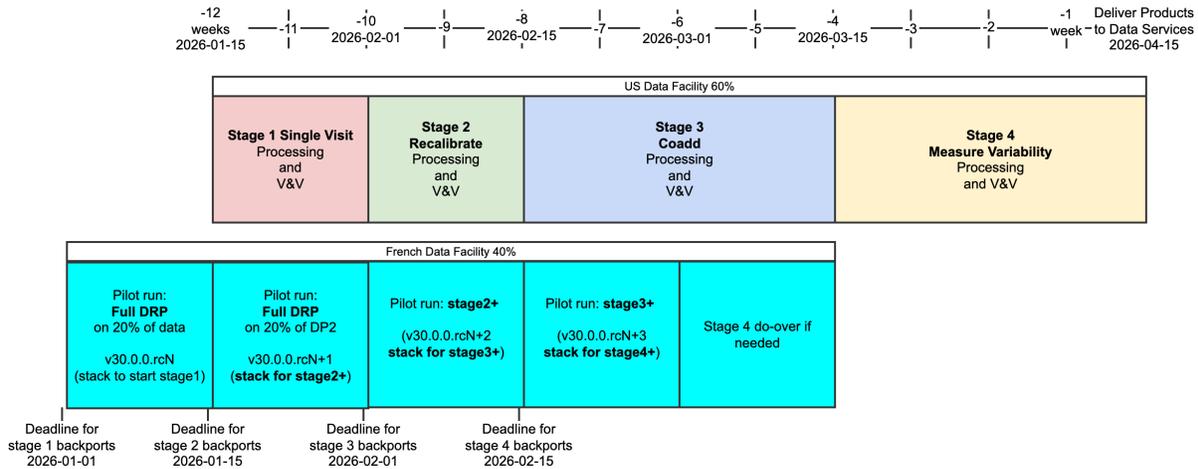


FIGURE 2: Planned timing and distribution of DP2 processing work across SLAC and IN2P3.

completeness. The Verification and Validation team is responsible for monitoring data quality throughout production and for signing off on the final data products at each stage.

A Data Release Steering Committee will coordinate decision-making across these groups. The committee will include one representative from each team and will approve decisions that span their shared responsibilities, including the resolution of quality issues, the incorporation of backports, the need for partial or full reprocessing, and other production decisions that do not require escalation to the project-level Data Release Board.

3.1 Data Production Timeline

For reference by the Data Production division, these are the internal milestones for DP2 Production as of early December 2025. These are not user-facing delivery dates. The user facing delivery date for DP2 is July-Sept 2026.

- 2025-12-03: Feature freeze for the DP2 processing software. All new functionality must be merged by this date. Subsequent changes are limited to approved backports.
- 2026-01-02: Final full pilot run begins. This pilot run will exercise the complete processing configuration and will be used to validate readiness for production.
- 2026-01-16: Final pre-run pilot is delivered to Data Services with a shepherd. This deliv-

ery is used to validate data access patterns, service interfaces, and end-to-end compatibility. Stage 1 processing of DP2 begins.

- 2026-01-31: Stage 2 processing of DP2 begins.
- 2026-02-15: Stage 3 processing of DP2 begins.
- 2026-03-17: Stage 4 processing of DP2 begins. Templates required for DIA will be available by this date.
- 2026-04-16: Final DP2 data products are delivered to Data Services with a shepherd for integration into user-facing services.

A member of the Pipelines team will serve as the shepherd traveling with the data products (the sheep) delivered to Data Services.

4 Data Services

5 Planning

We have some milestones and epics leading up to DP2.

Label milestones and epics DP2 to have them appear

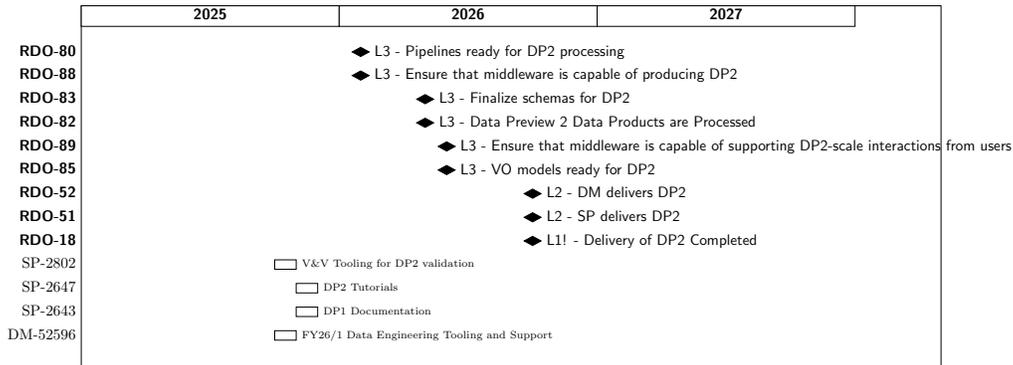


FIGURE 3: Data Preview 2 - plan

Open milestones are listed in Table 4.

Table 4: Milestones for Rubin Observatory Data Management and System Performance

Milestone	Jira ID	Rubin ID	Due Date	Level	Status	RubinTeam
L3 - Pipelines ready for DP2 processing	RDO-80	not set	2026-01-15	3	Proposed	Algorithms & Pipelines
L3 - Ensure that middleware is capable of producing DP2	RDO-88	not set	2026-01-31	3	Incomplete	Pipeline Middle-ware
L3 - Finalize schemas for DP2	RDO-83	not set	2026-04-01	3	Proposed	Data Engineer-ing
L3 - Data Preview 2 Data Products are Processed	RDO-82	not set	2026-04-16	3	Proposed	Campaign Man-agement
L3 - Ensure that middleware is capable of support-ing DP2-scale interactions from users	RDO-89	not set	2026-05-01	3	Proposed	Pipeline Middle-ware
L3 - VO models ready for DP2	RDO-85	not set	2026-05-01	3	Proposed	Data Engineer-ing
L2 - DM delivers DP2	RDO-52	not set	2026-09-30	2	Incomplete	RDM Manage-ment
L2 - SP delivers DP2	RDO-51	not set	2026-09-30	2	Incomplete	RPF Manage-ment
L1! - Delivery of DP2 Completed	RDO-18	not set	2026-09-30	1	Incomplete	Directorate

A Acknowledgements

This material is based upon work supported in part by the National Science Foundation through Cooperative Agreements AST-1258333 and AST-2241526 and Cooperative Support Agreements AST-1202910 and AST-2211468 managed by the Association of Universities for Research in Astronomy (AURA), and the Department of Energy under Contract No. DE-AC02-76SF00515 with the SLAC National Accelerator Laboratory managed by Stanford University. Additional Rubin Observatory funding comes from private donations, grants to universities, and in-kind support from LSST-DA Institutional Members.

B References

C Acronyms

Acronym	Description
AOS	Active Optics System
AST	NSF Division of Astronomical Sciences

AURA	Association of Universities for Research in Astronomy
COSMOS	Cosmic Evolution Survey
DDF	Deep Drilling Field
DE-AC02	Department of Energy contract number prefix
DESI	Dark Energy Spectroscopic Instrument
DIA	Difference Image Analysis
DM	Data Management
DP1	Data Preview 1
DP2	Data Preview 2
ECDFS	Extended Chandra Deep Field-South Survey
EDFS	Euclid Deep Field South
FGCM	Forward Global Calibration Method
FWHM	Full Width at Half-Maximum
GBDES	Gary Bernstein Dark Energy Survey
IN2P3	Institut National de Physique Nucléaire et de Physique des Particules
L1	Lens 1
L2	Lens 2
L3	Lens 3
LSST	Legacy Survey of Space and Time (formerly Large Synoptic Survey Telescope)
LSST-DA	LSST Discovery Alliance
LSSTCam	LSST Science Camera
LSSTComCam	Rubin Commissioning Camera
M1M3	Single piece of glass for Primary Mirror/Tertiary Mirror
PSF	Point Spread Function
RDM	Rubin Data Management
RDO	Rubin Directors Office
RPF	Rubin system PerFormance
RTN	Rubin Technical Note
SLAC	SLAC National Accelerator Laboratory
SP	Story Point
ToO	Target of Opportunity
US	United States
USDF	United States Data Facility

VO	Virtual Observatory
XMM	ESA X-ray Multi-mirror Mission

Draft