Rubin-Euclid Derived Data Products

Initial Recommendations

Leanne Guy & Jean-Charles Cuillandre (WG co-chairs)
on behalf of the DDP Working Group and Community

Rubin-Euclid Derived Data Products

Rubin & Euclid senior management appointed 13 scientists each to a Derived Data Products Working Group (DDP-WG) with a first specific charge 2021.

DDP-WG oversight committee :

- Rubin : Robert Blum, Phil Marshall, Željko Ivezić
- Euclid : Yannick Mellier, Jason Rhodes, René Laureijs

DDP-WG co-chairs : Leanne Guy for Rubin, Jean-Charles Cuillandre for Euclid

DDP-WG Rubin members :
Yusra Alsayyad, Etienne Bachelet, Manda Banerji, Franz Bauer, Jim Bosch, Tom Collett, Siegfried Eggl, Catherine Heymans, François Lanusse, Peter Melchior, Dara Norman, Michael Troxel

DDP-WG Euclid members :
Eric Aubourg, Hervé Aussel, Chris Conselice, Adriano Fontana, Henk Hoekstra, Isobel Hook, Konrad Kuijken, Joe Mohr, Michele Moresco, Reiko Nakajima, Stéphane Paltani, Daniel Stern

Charge to the Rubin-Euclid DDP-WG

- Focuses on the design not the creation of the proposed DDPs.
- The DDP-WG charter aimed to ensure both communities were consulted together, and at large, to gather input on the desired DDPs driven by science.
- The ongoing worldwide pandemic did not allow for the execution of the initial plan as laid out in the original charter.

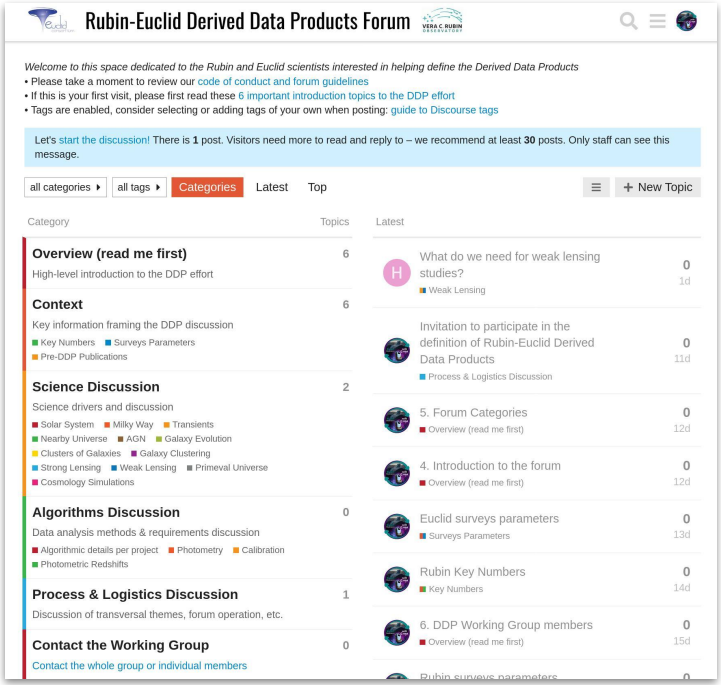
Charge: The DDP-WG should help plan a virtual or in-person workshop that is open to interested Rubin and Euclid data rights holders. This workshop should be focused on gathering community input into the desired initial DDPs. Based on the input from that meeting, the DDP-WG will:

- Design an initial set of DDPs that could be shared promptly and simultaneously with both the Euclid Consortium and the LSST Science Community for scientific use, in a way that protects the unique science of each collaboration and is consistent with both communities' data policies.
- Outline the scientific justification and quantify, approximately, its impact for each proposed DDP.
- **Issue an initial set of recommendations within 9 months** of the creation of the DDP-WG; these recommendations would be made to the ECB and Rubin Observatory Director.
- Set a cadence for virtual and in person meetings and workshops that they feel is consistent with developing recommendations for revised or new DDPs and then make those recommendations to the ECB and Rubin Observatory Director.
- Gather input from their respective communities about desired DDP.
- Focus only on designing DDPs, and not on issues of DDP creation or forming inter-project science collaborations.

DDP creation: The DDP-WG reports to and recommends DDPs to the ECB and Vera Rubin Observatory Director for approval. If approved, the respective consortia will then have to come to an eventual agreement about where, by whom, on what time scale, how, and with what funding the DDPs will be created.

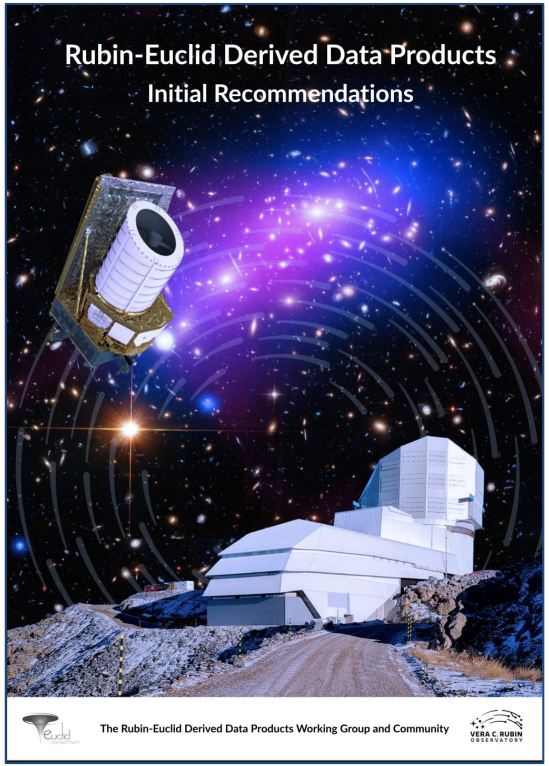
Model of Virtual Discussion

- Standalone, world-public, project-independent forum: https://community.rubin-euclid-ddp.org/
- Open to all data rights holders in both the Rubin Euclid communities – this approach ensured that everyone had a chance to contribute to the debate
- Categories for science topics and algorithms instigated by DDP-WG members (moderators) or other specialists from Rubin and Euclid
- Basis for the DDP-WG report.
- Contributions are mostly written exchanges but also include recording of talks and virtual sessions.
- 350 registered scientists.



The "Discourse - Civilized Discussion" platform ensures community exchanges with the power of asynchronicity across continents.

Rubin-Euclid DDPs – Initial Recommendations



- Report released 21 December 2021, 78 pages, 120 authors in total, uploaded to archive.
- Provides recommendations for an initial set of DDPs covering:
  - Solar System, Milky Way, Transients, Nearby Universe, AGN & Galaxy Evolution, Clusters of Galaxies, Galaxy Clustering, Strong Lensing, Weak Lensing, Primeval Universe
- Recommendations broadly grouped into two categories:
  - ***Cross-cutting DDPs (5):*** which will enable a wide range of complementary science goals,
  - ***Science-specific DDPs (58):*** which will enhance the science yield for a specific science case.
- Given the diversity in the complexity of the suggestions. We strongly advocate for a tiered approach to developing DDPs over the lifetime of both surveys: from simple catalog merging and cutout exchange, to full blown joint pixel processing.

Recommendations for cross-cutting Rubin-Euclid DDPs

DDP-1-CC: Multi-band Rubin+Euclid list-driven photometry catalogs:
Photometric redshifts are at the heart of the high-profile cosmology science cases of both surveys, with stringent accuracy requirements that cannot be met using a combination of two independent photometry catalogs. ***At a minimum, a*** list-driven photometry source exchange for point sources and galaxies detected in all r,i,z and Y,J,H bands across both catalogs above 5-sigma. **Timescale**: As soon as the two surveys overlap

DDP-2-CC: Multi-band Rubin+Euclid forced photometry catalog based on joint-pixel processing:
Starting with object detections across both surveys based on the DDP-1-CC selection function, measure PSF, aperture and total fluxes and/or upper limits across all bands using matched images in the other survey. **Timescale**: Post Euclid and Rubin DR1. Incrementally increasing in area/depth/complexity through the lifetime of both surveys

DDP-3-CC: Multi-band Rubin+Euclid deblended photometry catalog from joint-pixel processing:
Starting with object detections across both surveys based on the DDP-1-CC selection function, measure deblended component with VIS and total fluxes and/or upper limits across all u,g,r,i,z,y,Y,J,H bands using matched images in both survey datasets, while respecting the data policy driven DDP-1-CC source selection function. This represents the most complex approach. **Timescale**: Post Euclid and Rubin DR1. Incrementally increasing in area/depth/complexity through the lifetime of both surveys

Recommendations for cross-cutting Rubin-Euclid DDPs

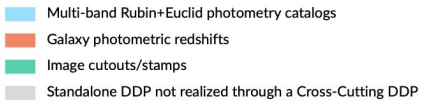
DDP-4-CC: Galaxy “pixel” photometric redshifts with machine learning:
Full probability distributions for the photometric redshift estimates are required for all science cases which need to propagate errors into physical parameters using a range of algorithms incorporating both empirical/training-set based methods and template-fitting run on the joint multi-wavelength catalogs. Joint-pixel analysis with machine learning will further benefit photometric redshift estimates at both surveys depth limits in particular when deblending becomes an issue for Rubin. Similar selection function as the above photometric catalogs DDPs.
Timescale: Post Euclid and Rubin DR1. Incrementally increasing in area/depth/complexity through the lifetime of both surveys

DDP-5-CC: Image cutouts/stamps delivery service:
Exchange of Image cutouts (pixels) on small areas of the sky will enable key scientific investigations, e.g transient science, strong lensing, and drop-out science. Sharing within the two projects a limited number of pixels driven by the angular size of the source of interest is compatible with the DDP definition considering the scientific return of a highly specific usage. Timescale: As soon as the two surveys overlap. **Timescale:** As soon as the surveys overlap

Tiered Prioritization of recommended DDPs

Cross-Cutting (CC)

DDP-1-CC	B	P1+U1+YR	T1	Multi-band Rubin+Euclid photometry list-driven catalogs
DDP-2-CC	B	P1+U2+DR	T2	Multi-band Rubin+Euclid forced photometry catalog from joint-pixel processing
DDP-3-CC	B	P2+U2+DR	T3	Multi-band Rubin+Euclid deblended photometry catalog from joint-pixel processing
DDP-4-CC	B	P2+U2+DR	T3	Galaxy "pixel" photometric redshifts
DDP-5-CC	B	P1+U1+RT	T0	Image cutouts/stamps delivery service



Column 1: DDP identifier

Column 2: Benefit, **B**=Both communities, **R**=Rubin, **E**=Euclid

Column 3	Tier	Description
P1 + U1 + RT	T0	Ready when both telescopes observe the same sky in 2023
P1 + U1 + YR	T1	In conjunction with the Rubin-LSST Year 1 release in 2025
P1 + U2 + DR	T2	In conjunction with the Euclid DR2 and LSST Year 3 in 2027
P2 + U2 + DR	T3	In conjunction with the Euclid DR3 and LSST Year 4 in 2029
Non-baseline	T4	Pending definition of Euclid's non-allocated time (illustrative DDPs)

P: Relative scientific **P**riority

U: **U**rgency, timescale on which the DDP is needed, e.g time-sensitive/small area DDPs might be worth producing and sharing before producing DDPs from a fuller analysis of a larger area.

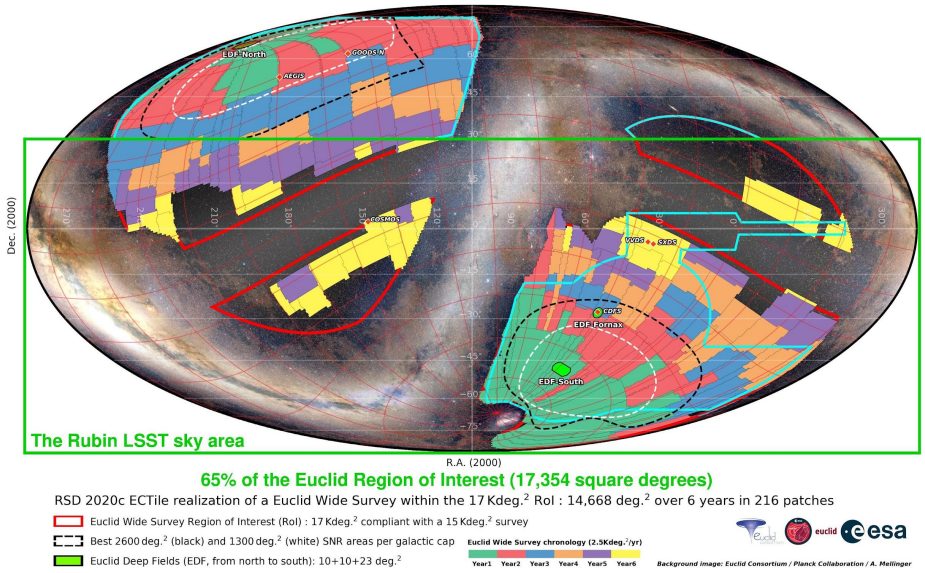
RT/YR/DR: Cadence for producing DDPs. RT= Real-Time (~ day) for transients, YR = "Yearly Release" matching the Rubin-LSST releases, DR = "Data Releases" for products that can wait for longer timescales, such as Euclid DR3

Differential Chromatic Refraction (DCR) model improvements

- Rubin's model for correcting differential chromatic refraction can be improved by a factor of ≈ 2 by incorporating pixel-level morphological information from Euclid's high resolution VIS band, which would drive improvements in essentially all downstream DDPs.
- The DCR model can only be improved over the region where the VIS data are available and is fundamentally rooted in joint-pixel level processing.
- Work will be required to understand how to incorporate Euclid morphological and spectral input into the Rubin DCR correction model. We recommend this to be investigated initially over the EDF-South DDF as a demonstration test case and, if a significant improvement is demonstrated, ultimately applied to the entire overlapping wide area in the long term.

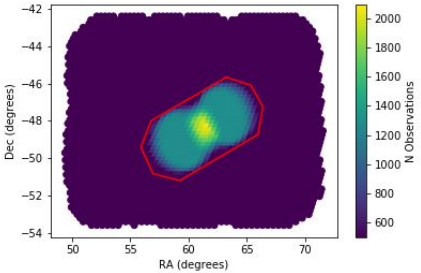
Survey Strategy Optimization for DDPs

A key factor in maximizing the impact of Rubin-Euclid DDPs is the coordination of each survey's observing strategy; maximizing the spatial and temporal overlap of the two surveys will enhance almost all science domains.

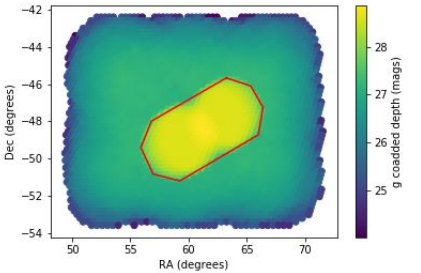


- Large overlap area of up to approximately 9000 square degrees
- 2 overlapping DDFs: Rubin and Euclid will both observe EDF-Fornax (10 sq deg) field. EDF-South (23 sq deg) is proposed as a Rubin DDF.
- LSST Cadence Note – Enhancing LSST Science with Euclid synergy and a mini-survey of the northern sky to Dec < +30 : modifications to the Rubin WFD towards an extended footprint driven by dust extinction limits to enhance Euclid synergy with up to a 9400 deg2 overlap.

Open sharing of data across small area of sky



Two Rubin Deep Drilling Fields at half-depth on EDF-South



... or a uniform depth on EDF-South with an optimal Rubin dithering

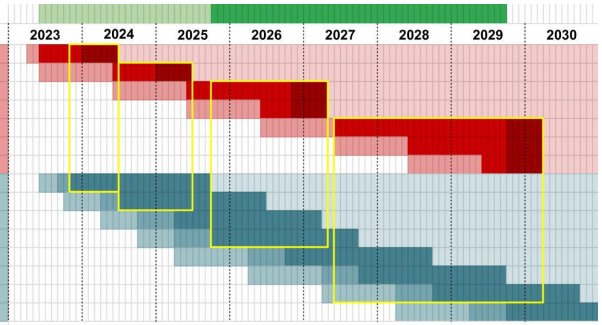
- Openly sharing all imaging pixel data over a common small area of the sky across both projects will enable the early development of methods, algorithms and software that will be beneficial to all DDPs before embarking on full scale JSP
- …..while also allowing both communities to investigate delivered DDPs in depths and propose paths for further improvements in future releases.
- We recommend for this purpose the proposed Rubin DDF over the 23 sq degree EDF-South.

Credit: P. Joachim

Timescales and Data Releases

Rubin-Euclid Coordination Timeline

Survey	Data Release	
Euclid	Q1	Misc. sky areas (EDF, etc), total 50 sq deg
Euclid	DR1	Euclid Y1 (2500 sq deg, << 1000 sq deg overlap)
Euclid	Q2	Euclid Y2
Euclid	DR2	Euclid Y3 (7500 sq deg, ~3000 sq deg overlap)
Euclid	Q3	Euclid Y4
Euclid	Q4 (TBC)	Euclid Y5
Euclid	DR3	Euclid Y6 (15000 sq deg, ~7000 sq deg overlap)
LSST	DP1	LSST ComCam
LSST	DP2	LSST SV (~1000 sq deg, 180 visits / Y2 depth)
LSST	DR1	LSST First 6 Months
LSST	DR2	LSST Y1 (90 visits)
LSST	DR3	LSST Y2 (180 visits)
LSST	DR4	LSST Y3 (270 visits)
LSST	DR5	LSST Y4 (360 visits)
LSST	DR6	LSST Y5 (450 visits)



Assumptions:

February 2023	Euclid mission launch date
April 2024	LSST survey start

Surveys color keys:

	Observing
	Processing
	Proprietary Access
	Public Access

Notes:

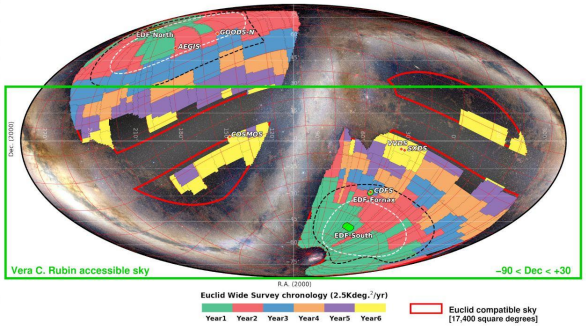
(a) LSST data release dates may move by +/3 months as the operations team adapts to circumstances.

(b) Euclid plan additional quick releases containing specific featured data products made with the Y2 ("Q2") , Y4 ("Q3") and Y5 ("Q4", TBC) data. The data from these years will be available to the Euclid Consortium to use while they are being processed, there just won't be an internal release of a full data release dataset.

(c) The overlap between Euclid Y1 and LSST SV is potentially quite small, because Rubin commissioning observations are needed at a wide range of latitude (and the best calibration pre-cursor data tends to be closer to equatorial). The SIT-Com team's field selection is not yet determined.

(d) DDP transient science can start in 2023 with limited sky overlap (green bar).

(e) LSST Y1 leads to matched survey depths for photo-z estimation: the production of related DDPs (photometric catalogs) spans 4 years (top darker green bar, 2025 to 2029) based on LSST Y1 to Y4 yearly data releases progressively matched to the Euclid survey increasing overlap.



Credit: Phil Marshall & Jean-Charles Cuillandre

- Recommended DDPs should be produced as soon as Rubin and Euclid observe a common area of the sky, e.g as early as 2023 in an investigative manner (limited overlap) in order to be mature in time for 2025: Euclid DR2 + LSST Y1
- DDPs that will enable transient science should materialize on a short timescale, e.g. 24 hr and based on a fast joint processing.
- More complex DDPs would fit better in the context of the annual Rubin data release scenario

Next steps

- The work of the Rubin-Euclid DDP Working Group is complete, the report is issued.
- If the report recommendations are approved, the respective consortia can then proceed to the funding and implementation phase; decisions about where, whom, on what timescale, with what funding, and how the recommended DDPs are produced
- As both projects approach first light, we recommend the prompt creation of working groups and task forces charged with producing the DDPs recommended in Rubin-Euclid DDP-WG report. This should include:
    - a task force to explore the detailed software and computing requirements,
    - a joint simulations group, and
    - a joint implementation group to realize the recommendations of this report.
- The Community forum will remain open and can be used for ongoing discussion into the implementation phases.