Euclid/Rubin Derived Data Product Implementation Working Group

Eric JULLO (LAM)

On behalf of the DDP-IWG

What are the DDP?

- Jointly derived data products (DDPs) are constructed from the combination of Rubin and Euclid data, and have the potential to open up new discovery space beyond that of what each mission can achieve independently.

Context – The DDP-WG

Rubin and Euclid management appointed in 2020, 13 scientists each to a Derived Data Products Working Group (DDP-WG)

- Rubin members (13):

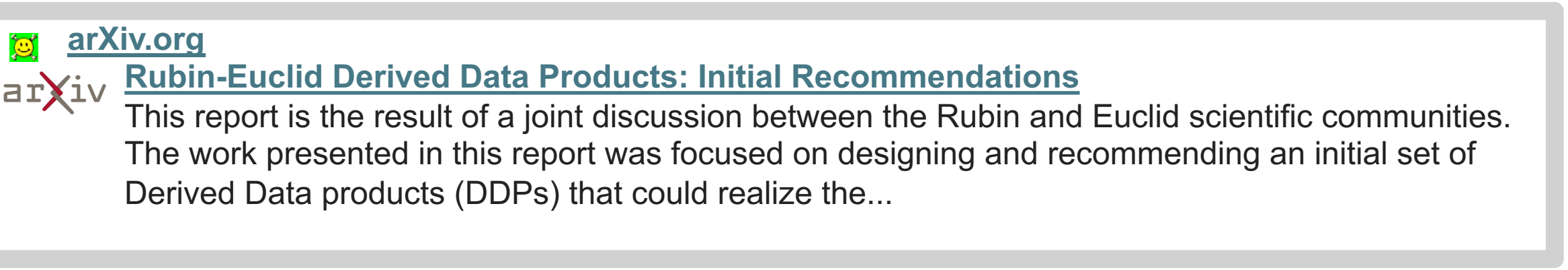
***Leanne Guy**, Yusra Alsayyad, Etienne Bachelet, Manda Banerji, Franz Bauer, Jim Bosch, Tom Collett, Siegfried Eggl, Catherine Heymans, François Lanusse, Peter Melchior, Dara Norman, Michael Troxel*

- Euclid members (13):

***Jean-Charles Cuillandre**, Eric Aubourg, Hervé Aussel, Chris Conselice, Adriano Fontana, Henk Hoekstra, Isobel Hook, Konrad Kuijken, Joe Mohr, Michele Moresco, Reiko Nakajima, Stéphane Paltani, Daniel Stern*

The DDP-WG report

- Initial recommandations from the DDP-WG (12/2021)



78 pages, 120 co-authors:

- the first 14 pages gives an overview.
- 10 independent science sections + survey optimisation discussion.
- 1 page executive summary for those short in time

« All interested Rubin and Euclid data rights holders were invited to contribute via an online discussion forum and a series of virtual meetings. »

=> https://community.rubin-euclid-ddp.org

The DDP-WG report content

- 63 DDP identified with each 5 key aspects:
  - Science Case, Nature of the Derived Data Products, Algorithms, Computing, Timescale
- 10 Topics:
  - Cross-Cutting, Solar System, Local Volume, Galactic Plane, Local Universe, Transients, Galaxy Evolution, AGN, Static Cosmology (WL, GC, Cl), Strong Lensing, Primaeval Universe

Illustration of the nomenclature:

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

DDP ***Implementation*** WG

Rubin and Euclid management appointed in Dec 2023, 14 scientists to an Implementation Derived Data Products Working Group (DDP-IWG)

- Rubin members (8):

Tod Lauer *(NoirLab), Siegfried Eggl (Illinois), Catherine Heymans (Edinburg), Peter Melchior (Princeton), Agnès Ferte (SLAC), YuanYuan Zhang (NoirLab), Aaron Meisner (NoirLab), James Chiang (SLAC)*

- Euclid members (6):

James Colbert *(IPAC), Joe Mohr (LMU), Guillaume Libet (CNES), Eric Jullo (LAM), Eduardo Balbinot (RUG), William Hartley (UNIGE)*

DDP IWG - Charges

- Establish a roadmap for the creation of the 63 DDPs, explicitly identifying which DDPs can and should be created immediately (i.e. without significant new resources).
- Identify what technical capabilities individuals need to have to develop the 63 DDPs and, where possible, identify candidates with those capabilities and skills.
- Identify infrastructure requirements for the creation of the DDPs.
- Identify which bodies and organizations might financially support the creation of the DDPs and develop cases to obtain that support.
- Deliver a roadmap to the Rubin and Euclid leadership and continue to update Rubin and Euclid leadership on a quarterly basis on its progress.

DDP IWG - Meetings

- Group meetings every two weeks
- Minutes available on https://community.rubin-euclid-ddp.org
- Level of secrecy vs transparency
  - Previous experiences in the DDP-WG led to unpleasant situations
  - Release of meeting notes on a best effort bases after approval from the group here https://community.rubin-euclid-ddp.org/c/iwg-meeting-notes-released . Latest on April 11th.

Community survey

- Distributed in many ways (Emails, Slack, Rubin Community Forum, etc)

- List of points
  - Group description
  - DDPs to which will contribute
  - Rubin/Euclid data access
  - Level of effort
  - DDP creation challenge
  - Other info desired

- Deadline May 8th … but survey is still open

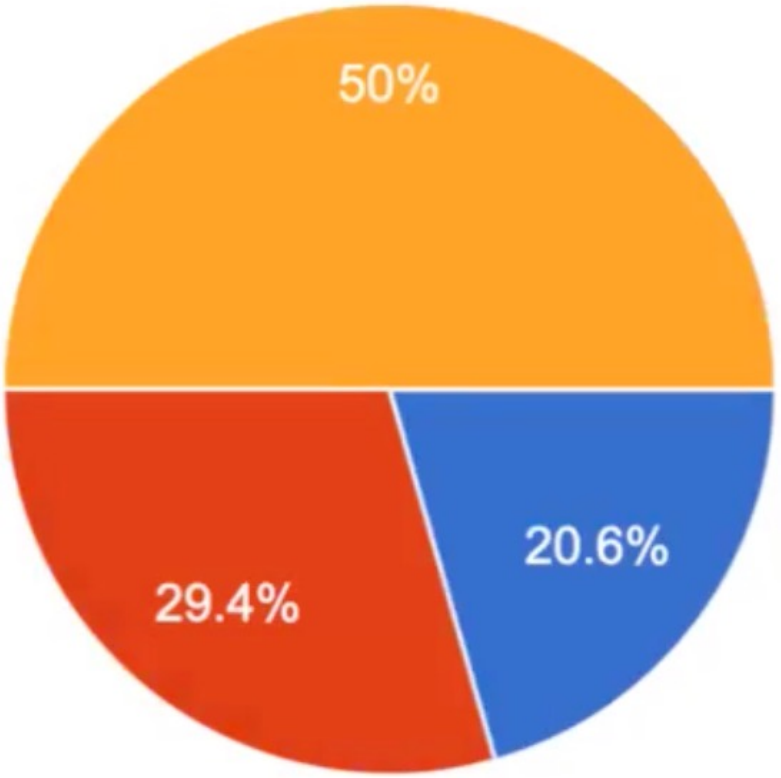
Community survey preliminary outcomes

- The survey is still being digested by the DDP-IWG
- 34 responses (as of 1pm PDT, May 8), 8 on the French side:

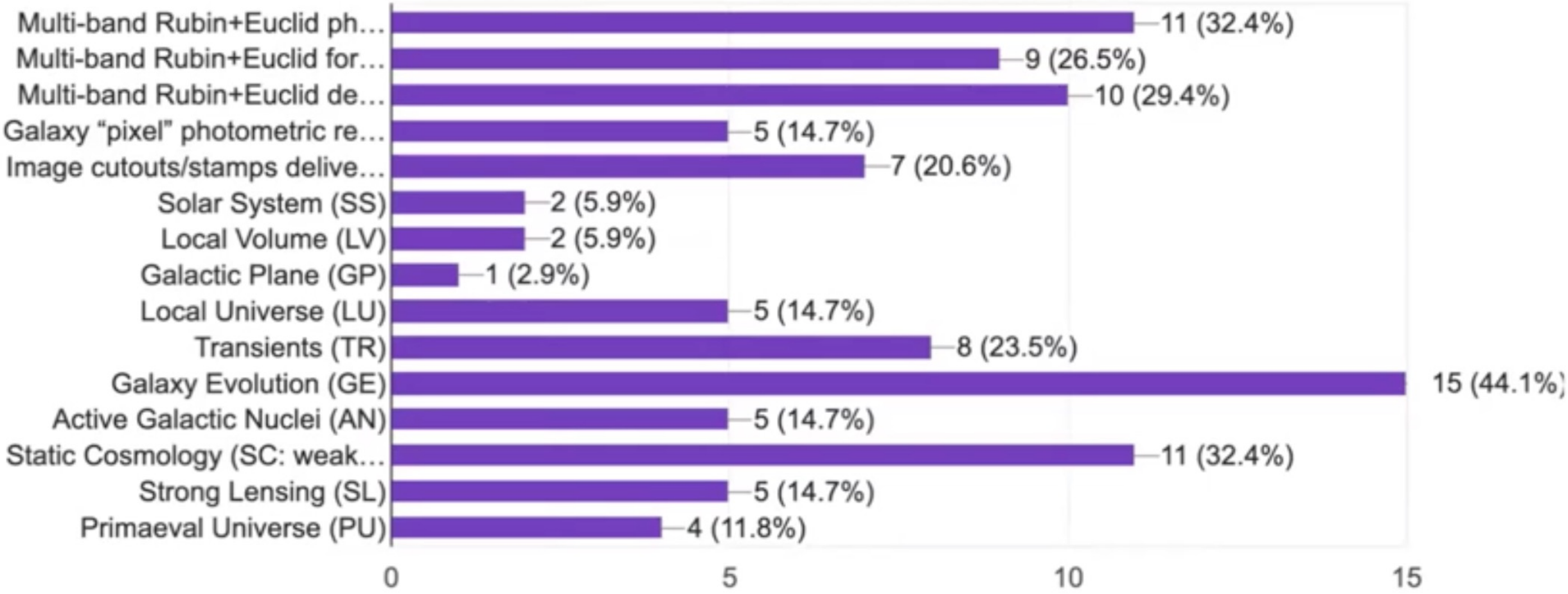
  P.-A. Duc (UniStra), M. Ricci (APC), S. Mei (APC), E. Aubourg (APC), S. Dagoret-Campagne (IJCLab), B. Carry (OCA), R. Gavazzi (LAM), A. Pisani (CPPM)

- Most of the groups require a roadmap, resources, etc (ie. the expected outcome of the DDP-IWG 😉)
- Agnès Ferté volunteered to dig through the survey, track names, and connect them to DDPs

Groups Data access

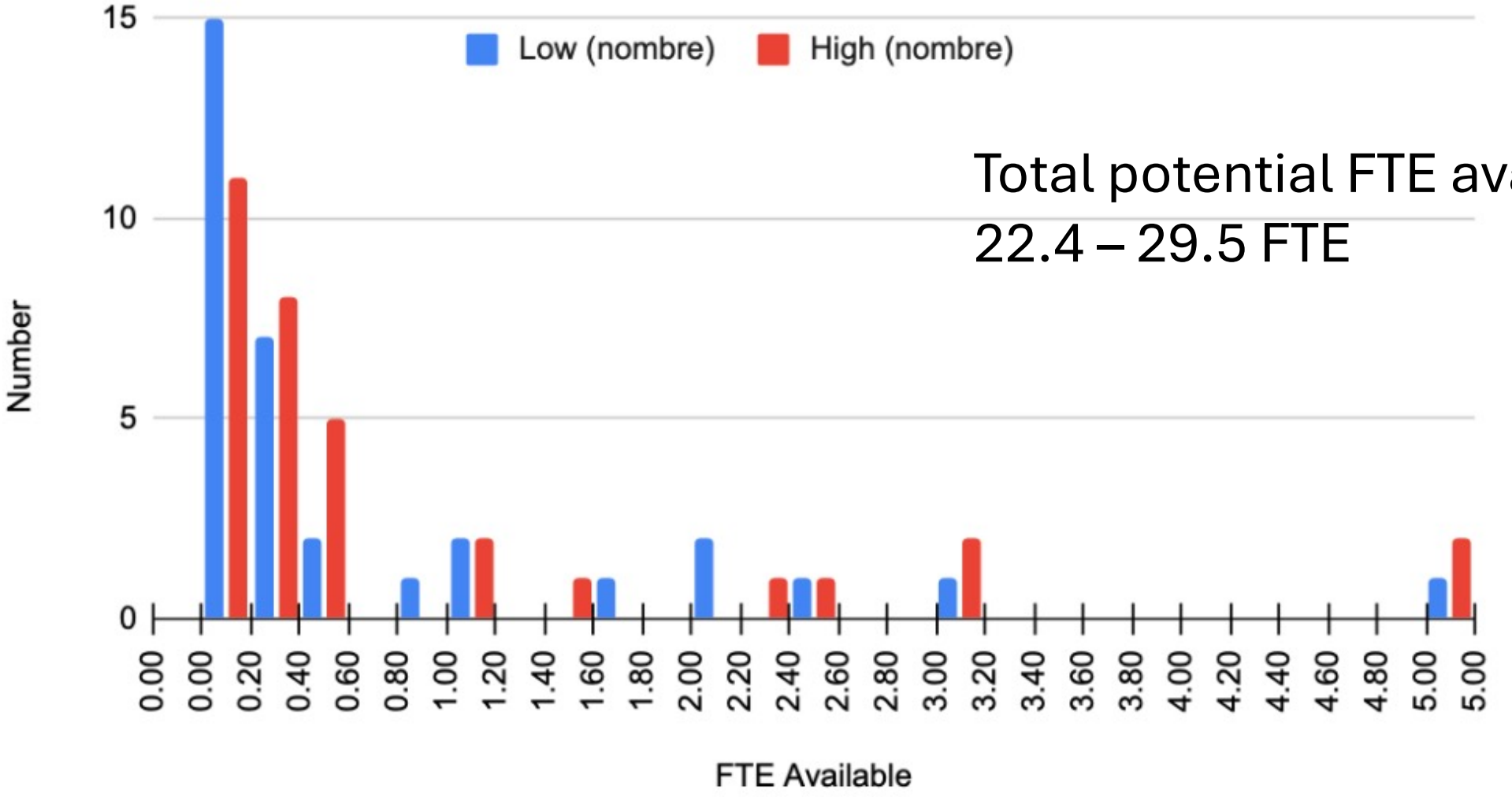


DDP that you/your group wish to contribue R&D effort?



=> Most of the categories are covered

Level of effort (FTE)



Total potential FTE availabe:
22.4 – 29.5 FTE

2/ Data sharing

⇒ Access rights: acknowledgement that the data is proprietary, with a statement of consequences if rules are broken.

⇒ Protection of proprietary data: Rubin's depth & Euclid spatial resolution and NIR data. Want to avoid the use of **low-SNR** detections to get Rubin photometry or Euclid morphology or red colors

⇒ Still we want to get the best out of the 2 datasets => deblend using all possible information and distribute according to agreed-upon rules.

2/ Discussion about implementation

Still some technical issues: Euclid Rate-limited cutout service

=> Not designed to be massive, just for validation checks. Would need to be performed on a one-time delivery service, based on a list of objects

Processing resources on the Rubin and Euclid side :

$\Rightarrow$ Rubin NSF/ACCESS program, with yearly renewed allocation

$\Rightarrow$ Euclid/Rubin CC-IN2P3 allocation

2/ Discussion about implementation

- Highest priority

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-2 & DDP-3 will produce a single, optimally produced joint catalog without restriction. We would then apply filtering in order to meet proprietary restrictions. DDP-2 might differ from DDP-3 in the restrictions we apply
- DDP-1 reason of existence is speed at which it can be accomplished. Need to define the format of the catalog

3/ Letters of support for funding

- US proposal for joint Roman/Euclid/LSST simulations PI: J. Rhodes
- UK proposal on transient/SN PI: M. Vincenzi
- JPL machine learning group proposal PI: U. Rebbapragada
- OpenUniverse simulation proposal PI: R. Mandelbaum, J. Rhodes

4/ Simulations

- Simulation subgroup lead: Tod
- Processing: no processing resources in the US. Only marginal at SLAC. Most likely being processed at CC-IN2P3 => advantage of having Rubin & Euclid data in the same place
- Data sharing issue => Same as for real data
- Data format issue => Proceed by increasing complexity
  - 1/ Produce SingleVisit (30s) with Euclid OU-SIM pipeline & reduce / stack with Euclidized Rubin pipeline
  - 2/ Process OU-SIM SingleVisit products with Rubin pipeline
  - 3/ Adapt LSST input galaxy catalogs to Euclid format & vice-versa

4/ Simulations – Current status

- LSST & Euclid simulated data on the COSMOS field :
  - SingleVisit Euclid-WIDE depth (15 x 30s exposures in 5 LSST bands)
  - Reduced & stacked with the Euclidized Rubin pipeline

5/ DDP timeline

- September 2024 - High level roadmap completed
- March 2025 - Euclid QR1
- October 2025 - Rubin starts main survey
- February 2026 - Final IWG Report, includes design spec document for each DDP
- March 2026 - Euclid DR1
- January 2031 – Final public delivery of Euclid 6-year mission data