Rubin-LSST@Italy
in-kind program status
& ICT perspectives

October 18, 2024

M. Brescia

UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II

Rubin is now a real telescope!
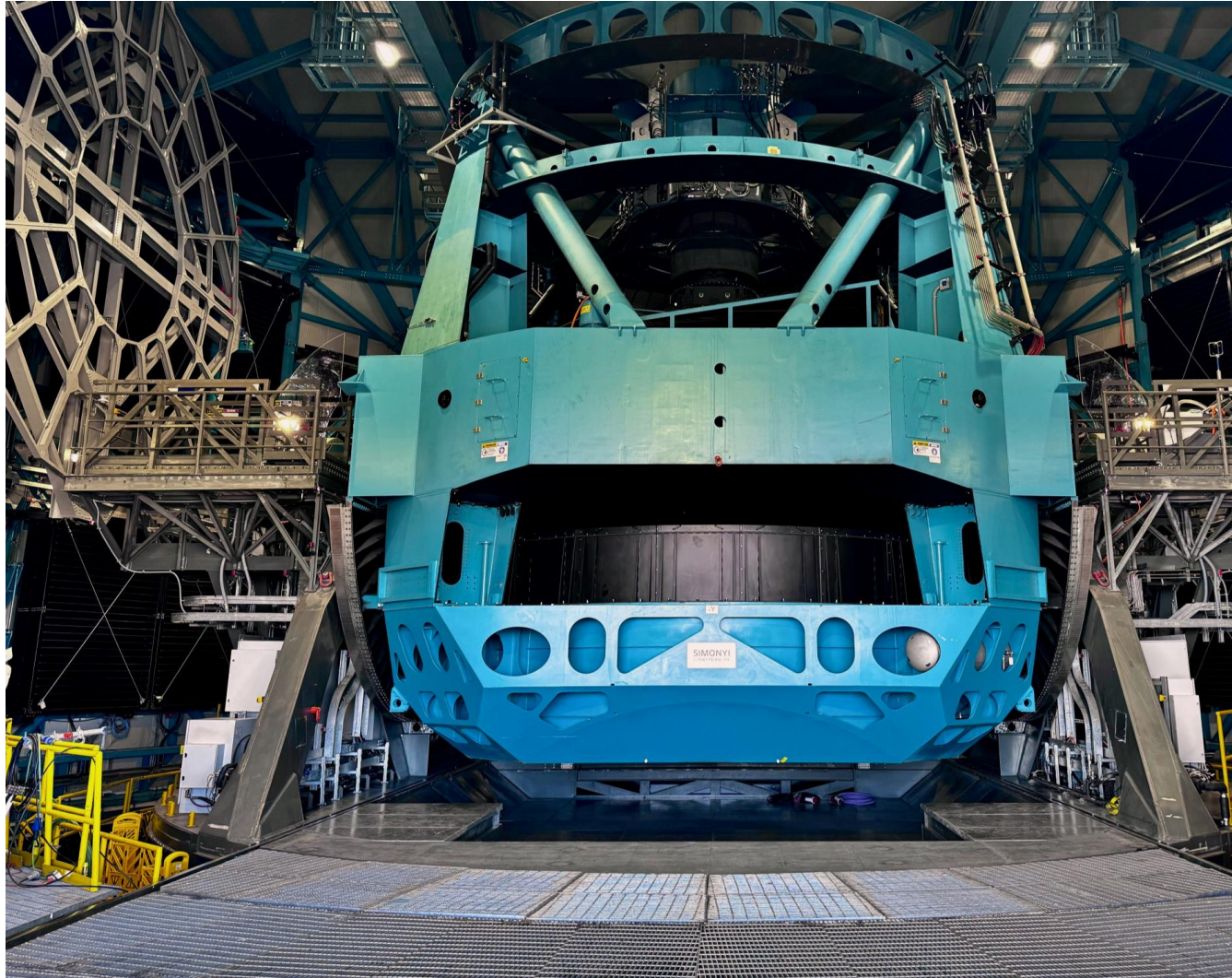
Two weeks ago, Rubin Observatory's M1/M3 mirror system was installed.

With all three mirrors and the commissioning camera now in place, we officially, for the first time in 15 years of construction, have a complete telescope!

INAF consistently contributed with commissioning on-site group thanks to the in-kind program
(**G. Rodeghiero, R. Canestrari, E. Giro, L. Rosignoli, F. Cusano**)
Rodeghiero et al. 2024

Next step: get this telescope on-sky!

Rubin LSST Data Rate

What LSST will Deliver:
A Data Stream, a Database, and a (small) Cloud

Level 1
- A stream of ~10 million time-domain events per night, detected and transmitted to event distribution networks within 60 seconds of observation.
- A catalog of orbits for ~6 million bodies in the Solar System.

Level 2
- A catalog of ~37 billion objects (20B galaxies, 17B stars), ~7 trillion single-epoch detections ("sources"), and ~30 trillion forced sources, produced annually, accessible through online databases.
- Deep co-added images.

Level 3
- Services and computing resources at the Data Access Centers to enable user-specified custom processing and analysis.
- Software and APIs enabling development of analysis codes.

Data Rights are crucial for the Italian scientific community

Italian participation to Rubin-LSST – in-kind program

The Italian scientific and technological community is heavily involved since 2016 in Rubin-LSST project, covering leadership roles in various WPs of the Science Collaborations (SC): R&D in scientific pipelines, scientific roadmaps and observational strategies, commissioning on-site...
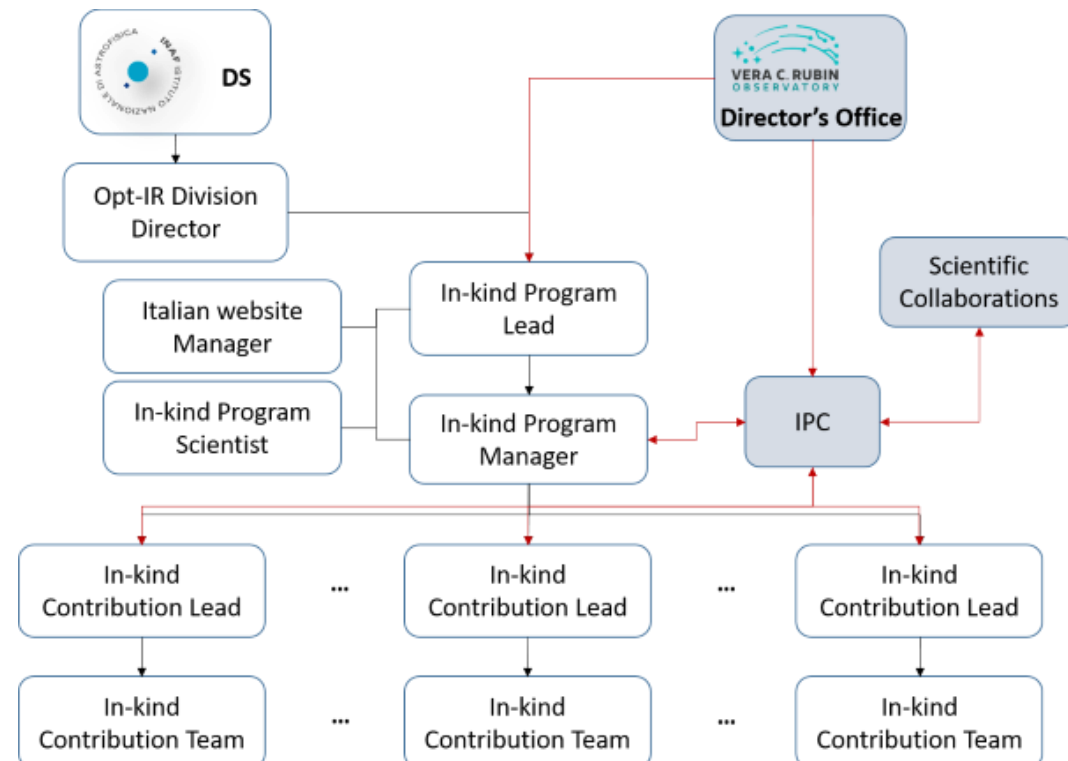
in-kind contribution categories:

- **Directable SW**: science SW for SCs for LSST pipeline
- **Commissioning**: on-site support & SW solutions
- **Observing Facilities**: guaranteed observing time & science products from observing programs with national instruments
- **External data**: photometric and/or spectroscopic data provided in exclusive preview
- **Non directable SW**: SW solutions already supplied and scientifically validated for SCs
- **DAC**: Data Access Centers, i.e. computing resources or science data production services

The Italian participation in the Rubin-LSST project is focused on the in-kind contributions program, for which the LSST Corporation undertakes to provide shares of data access rights, proportional and commensurate with the value of the In-kind contributions

More information on the official site:

https://sites.google.com/inaf.it/rubin-lsst-italy

Rubin-LSST Italia – in-kind program status

M. Brescia

Program Manager: drs. C.M. Raiteri

Data rights value: ~ 28%

In-kind contributions already started (2021 – 2023):

Commissioning

S22 **Contribution Lead**: Giuliana Fiorentino (start date: October 2021)
Staff effort in support of Rubin commissioning: Image quality analysis

Non directable

S25 **Contribution Lead**: Leo Girardi (start date: October 2022)
Population models of the LSST stellar content

Directable

S2 **Contribution Lead**: Angela Bongiorno (start date: October 2022)
Simulations of high-z AGNs and galaxies in the LSST survey

Commissioning

S21 **Contribution Lead**: Gabriele Rodeghiero (start date: October 2022)
Staff effort in support of commissioning: Engineering support at Telescope commissioning

Commissioning

S23 **Contribution Lead**: Giuseppe Riccio (start date: October 2022)
Support of Rubin commissioning: ML tools for instrumental monitoring and analysis

Rubin-LSST Italia – in-kind program status

Data rights value: ~ 16%

In-kind contributions started in 2024:

Directable → S6 **Contribution Lead**: Michele Cantiello
Tools for the measurement of surface brightness fluctuations on LSST data

Directable → S9 **Contribution Lead**: Crescenzo Tortora
Structural parameters with Machine learning

Directable → S10 **Contribution Lead**: Massimo Dall'Ora
Software Tools for Stellar Populations in Crowded Fields

Directable → S11 **Contribution Lead**: Laura Inno
Advanced active objects' detection & characterization

Directable → S15 **Contribution Lead**: Gisella Clementini
Tools for classification, full characterization and validation of variable sources

Rubin-LSST Italia – in-kind program status

M. Brescia

VST/LBT In-kind contributions starting in 2025:   **Data rights value: ~ 29%**

Obs. Facilities →

S1.1 **Contribution Lead**: Maria Teresa Botticella
A VST survey to support the Legacy Survey of Space and Time

Obs. Facilities →

S1.2 **Contribution Lead**: Claudio Grillo
Exploiting the synergy between LSST and VST to investigate the cosmos: the expansion rate and the geometry of the Universe measured through the time delays of strongly lensed variable sources

Obs. Facilities →

S1.3 **Contribution Lead**: Maurizio Paolillo
Toward next-generation time-domain surveys (TIMEDOMES): VST monitoring of the LSST Deep Drilling Fields

Obs. Facilities →

S18 **Contribution Lead**: Enrica Iodice
Access to Telescope Time for the US/Chilean communities

A workshop dedicated to VST/LBT In-kind contributions is scheduled in February 2025 with all involved LSST Science Collaboration representatives

Rubin-LSST Italia – in-kind program status

M. Brescia

In-kind contributions starting in the next future:

Data rights value: ~ 24%

Directable → S3 **Contribution Lead**: Rosaria Bonito
Development of a SW to classify variable stars

Directable → S5 **Contribution Lead**: Innocenza Busà
Software for the measure of the optical fluxes from galactic diffuse medium

Directable → S8 **Contribution Lead**: Ilaria Musella
Tools for the simulation of Pulsating Stars

Directable → S12 **Contribution Lead**: Sara Lucatello
Machine Learning tools for the characterization of stellar populations

Directable → S13 **Contribution Lead**: Emiliano Merlin
Advanced tools for extragalactic photometry

Directable → S14 **Contribution Lead**: Alex Saro
Tools for the simulations, detection, and characterization of galaxy clusters

Non directable → S16 **Contribution Lead**: Sergio Campana
The Son of X-Shooter contribution to Rubin: a set of 2,000 spectra for ML light curve classification

Non directable → S17 **Contribution Lead**: Piero Rosati
A Cluster Spectroscopy Hub for galaxy cluster science with LSST
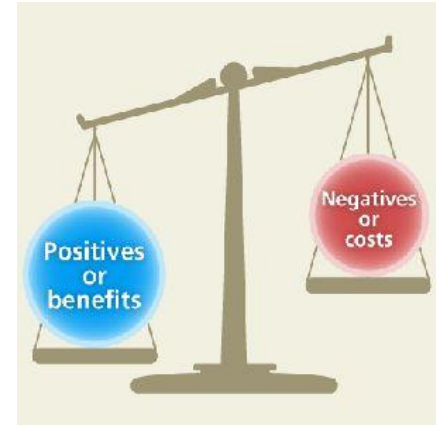
Non directable → S26 **Contribution Lead**: Angelo Antonelli
A Bridge from Gamma to Optical

Rubin-LSST Italia in-kind program numbers

Data Rights provide same real-time data access as US/Chilean community

INAF global investment in 15 years	
Missions/Meetings	430.000,00 €
AdR contracts	700.000,00 €
TD contracts	2.255.000,00 €
Contractor's HW	48.000,00 €
TOTAL	**3.433.000,00 €**

in-kind program duration: **15** years (*LSST survey duration: 10 years*)

interested national Researchers: 177 from INAF + 37 from other Institutes/Universities = **214**

in-kind contributions in 15 years: **23** *(+ IDAC)*

FTEs from Staff in 15 years: **51.5**

FTEs from contracts in 15 years: **48.8**

PI: senior (postdoc for > 10 years)

JA: junior (postdoc for ≤ 10 years)

Total contributed FTEs in 15 years: **100.3**

Total Data Rights (DR) for 10 years: **96 PIs + 384 JAs = 480** DR **(*IDAC proposal still under negotiation*)**

~715 € / (DR per year)   *past policy (15 quintuplets) cost before 2021: **4.000** € / (DR per year) for a total of **75 DR** and a total cost of 3.000.000 €, excluding the supplementary required budget for missions/meetings; Thus, the same cost but with a **net benefit of +85% for DR amount** (excluding IDAC return)*
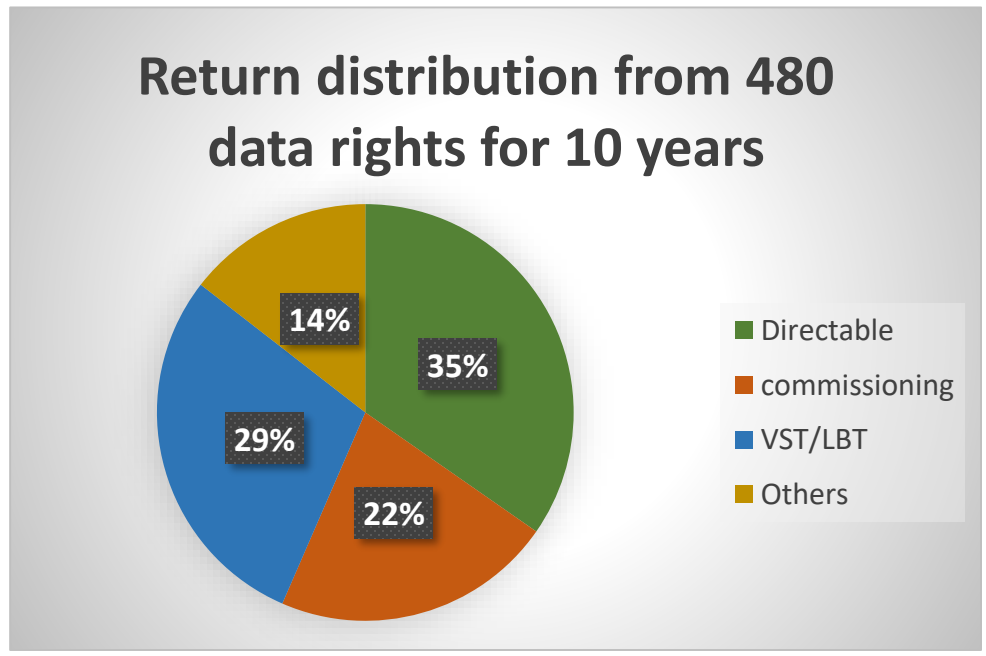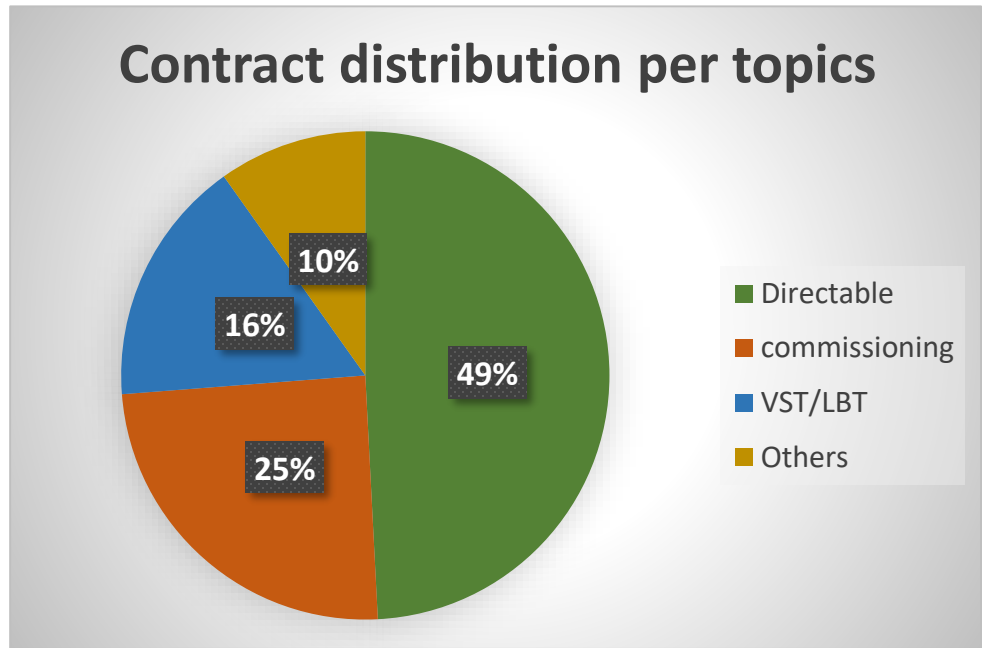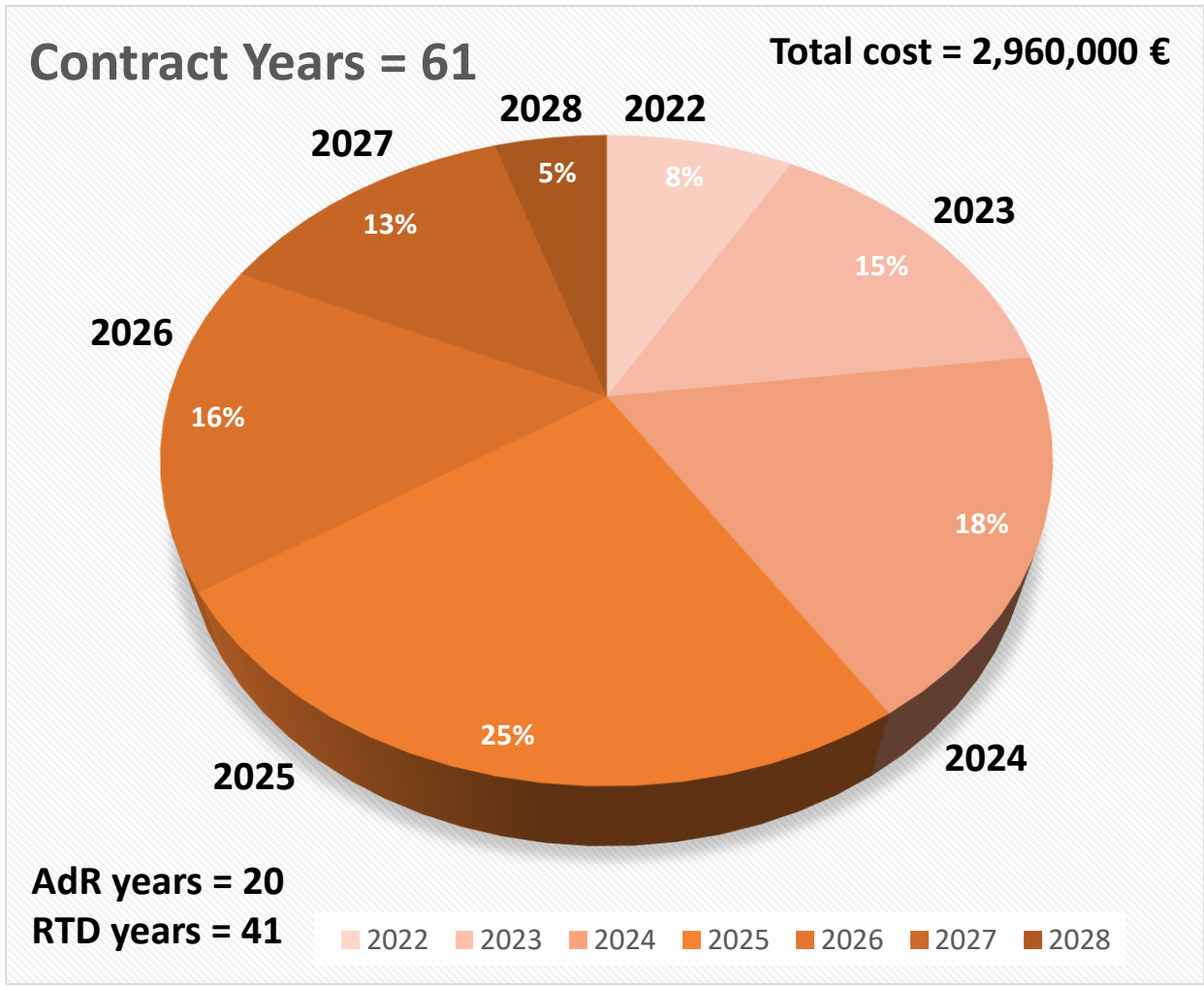
~34.000 € / FTE   *less than the standard cost for 1 FTE of an AdR annual contract (35 K€)*

~5 DR / FTE   *Each FTE invested will produce 1 PI + 4 JA DR positions (1 quintuplet)*

Rubin-LSST Italia in-kind program numbers

M. Brescia

Contract Years = 61

Total cost = 2,960,000 €

2022: 8%
2023: 15%
2024: 18%
2025: 25%
2026: 16%
2027: 13%
2028: 5%

AdR years = 20
RTD years = 41

Contract distribution per topics

- Directable: 49%
- commissioning: 25%
- VST/LBT: 16%
- Others: 10%

Return distribution from 480 data rights for 10 years

- Directable: 35%
- commissioning: 22%
- VST/LBT: 29%
- Others: 14%

Rubin-LSST Italia – in-kind program timeline

M. Brescia

	2021	2022	2023	2024	2025	2026	2027	2028	2029
Proposal informal acceptance & DR quotation	■								
first in-kinds started (VST+IDAC in stand-by)	■	■	■						
Data Rights Holder's policy and list delivery		■	■	🟥	🟩	🟩	🟩	🟩	🟩
Funding problems & new in-kinds suspended			■						
Internal program review		■	■						
Obtained increased and flexible fundings			■						
Urgent new in-kinds funded and started			■						
IPC's in-kind general review			■						
Internal in-kind program review			■						
IDAC use cases investigation & proposal				🟥	🟥				
VST/LBT based in-kinds proposal & acceptance				🟩	🟩				
Rubin-INAF (in)formal program agreement		■			🟧				
IDAC Infrastructure available					🟩				
End of dedicated AdR/TD contracts									🟩

... 2035

Rubin-LSST Italia – Data rights Policy

M. Brescia

Policies

INDEX	REASON
Level 0 – Highest priority (in-kind team CL, JA or LSST program membership)	
0.1	Contribution Lead (CL) of an in-kind team
0.2	DP0 delegate or Catalyst/LSST-related fellowship
Level 1-2 High priority (in-kind team membership or ex-quintuplet PI-ship)	
1	member of an in-kind team with exposed FTEs
2	PI of an ex-quintuplet still active in LSST data exploitation
Level 4-5 Medium priority (active in-kind team externals or ex-quintuplet members)	
4	Active external member of an in-kind team (not exposing FTEs) suggested by the related CL with motivations:
4.1	*Urgent/needed task to be carried out, which requires a direct data access (including LSST SC/WG membership imposed by its internal policy)*
4.2	*In-kind contribution started and under scheduled working time (active contribution)*
4.3	*Recognized concrete effort within the in-kind team*
5	Active member of an ex-quintuplet suggested by the related PI with motivations:
5.1	*urgent/needed task to be carried out, which requires a direct data access (including LSST SC/WG membership imposed by its internal policy)*
5.2	*recognized concrete effort within the in-kind team*
Level 6 – Low priority (active in LSST science but externals to in-kind or ex-quintuplet teams)	
6	Researcher interested to any LSST-related scientific task (external to any in-kind/quintuplet team):
6.1	*with active membership to any Rubin SC/WG*
6.2	*working to a scientific use case outside in-kind or ex-quintuplet topics and requiring a direct data access*
6.3	*supervisor of any PhD student involved in LSST data exploitation requiring a direct data access*
Level 7 – Lowest priority (general research/academy people interested in LSST Science)	
7	people interested at any level in the Italian participation to Rubin-LSST not satisfying any of the above priorities

PI data rights coverage: ~ 72% (availability: ~ -28%)
JA data rights coverage: 100% (availability: ~ +79%)

IDAC proposal may partially recover the PI needs.
Expected a final deficit of ~ -14%

[Rubin-LSST @ Italy - Italian Data Right Policy](#)

Italian Rubin-LSST in-kind program – Data right policy

Data Right Policy

The international Rubin-LSST in-kind program foresees that data rights are assigned by the Rubin IPC (In-kind Program Committee) as return of accepted in-kind contributions. This is one of the subjects of an official MoA (Memorandum of Agreement) between INAF and Rubin Foundations Board, still under negotiation. As well known, participation in the in-kind contribution program is on a voluntary basis. This implies that some groups of Italian researchers agree to expose their resources and efforts to provide a scientific/technological contribution to the Rubin-LSST community, obtaining in exchange a certain amount of data rights, partly donated to the Italian community interested in having access to the LSST data before the public releases, equalizing their right of access to the US and Chilean communities.

Although without an official signed MoA, the Italian in-kind program started in the last quarter of 2021, by following a contribution activation schedule, regulated by a temporary INAF-Rubin agreement. The Italian participation to the in-kind program is lead and financed by INAF. The financial support was regulated by an initial resolution of the INAF CdA in 2016, which allocated 3 million euros for a 10-years participation in the program based on quintuplets, predecessor of the current in-kind program, subsequently updated in 2023 to 3.5 million euros through an agreement between the INAF Scientific Director and the Director of UTG-I.

An updated data rights holder list is foreseen to be published by November 2024

[International Data Rights Holder List | Rubin Observatory](#)

Italian Community ICT needs for the Rubin-LSST

The italian community (214 researchers, mostly INAF) is involved in all science/technology topics of LSST.

A dedicated i-DAC (italian Data Access Center) would be extremely suitable to make the community able to pursuit its science goals, nested within the in-kind program, and to directly increase the amount of data rights holders (as outcome of i-DAC contribution efforts).

Among the wide variety of science use cases, related to the in-house LSST exploitation, we can mention:

1) Crowding source fields photometry, such as the analysis of the galactic plane (for sure the galactic bulge);

2) Classification and characterization of sources from images and catalogues (both in the time domain and at different scales, from solar system to large scale structures), including specialized methods (such as Machine/Deep Learning);

3) Synergies with other surveys for the combined exploitation of merged data, cross correlation with legacy multi-wavelength data, coordination of follow-up groups.

Conclusions: minimal benchmarks

Obviously, all these use cases must be considered as concurrent resource demands, i.e. as activities to carry out in parallel.

A rough estimate of required dedicated storage is an incremental scenario, starting from ~3 PB

- Crowding source fields photometry, such as the analysis of the galactic plane (for sure the galactic bulge);

  BENCHMARK   **~100 GB of RAM for a single CCD image processing (for each core)**
  ~20 TFlops for each single mosaic image processing

- Classification and characterization of sources from images and catalogues (both in the time domain and at different scales, from solar system to large scale structures), including specialized methods (such as Machine/Deep Learning);

  BENCHMARK   **Multi-GPU (many-core) parallel computing strategy with ~80 GB of dedicated RAM, for training on ~100K image cutouts with deep learning models**

- Synergies with other surveys for the combined exploitation of merged data, cross correlation with legacy multi-wavelength data, coordination of follow-up groups. Here the most demanding operation is the cross-matching, depending on the amount of entries and column size of the matched tables.

  BENCHMARK   **Multi-core strategy, 0.3 PFlops required for cross-matching between 2PB sized (150-columns) catalogues**

M. Brescia
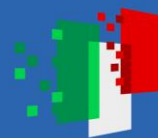
Nr. of CPU's cores requested
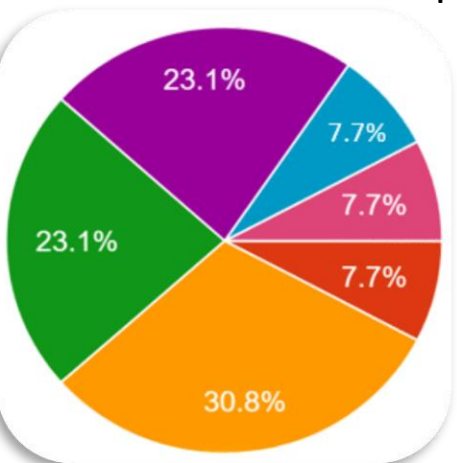
- from 1 to 4
- from 5 to 8
- from 9 to 16
- from 17 to 32
- from 33 to 64
- from 65 to 96
- more than 96

Primary projects involved

- ELT
- Euclid
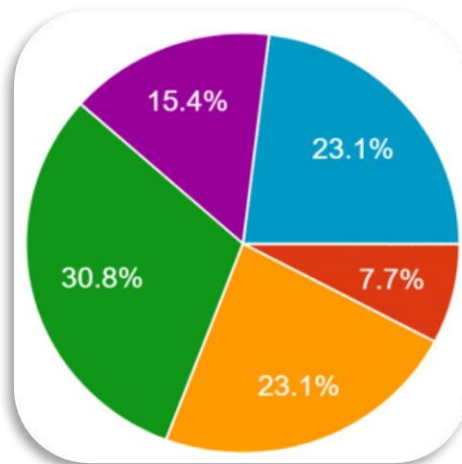- Rubin-LSST
- SKA
- VST
- Others

Nr. of GPUs requested

- no GPU required
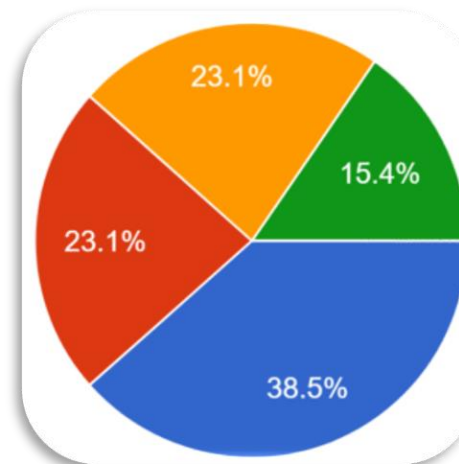- up to 1 GPU
- up to 2 GPUs
- up to 4 GPUs
- up to 6 GPUs

User Survey (15 research groups)

RAM requested

- up to 32 GB
- up to 64 GB
- up to 128 GB
- up to 256 GB
- up to 512 GB
- up to 1 TB
- up to 1.5 TB
- more than 1.5 TB

Storage requested

- up to 10 TB
- up to 20 TB
- up to 50 TB
- up to 100 TB
- up to 500 TB
- up to 1 PB
- up to 2 PB
- up to 4 PB
- more than 4 PB

GPU's RAM requested

- up to 16 GB
- up to 32 GB
- up to 48 GB
- up to 80 GB

UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II · Università degli Studi di Palermo · INAF ISTITUTO NAZIONALE DI ASTROFISICA NATIONAL INSTITUTE FOR ASTROPHYSICS · SAPIENZA UNIVERSITÀ DI ROMA

PNRR STILES – IR0000034, CUP C33C22000640006
Missione 4 – Componente 2 – Investimento 3.1

Missione 4 • **Istruzione e Ricerca**

take home message

M. Brescia
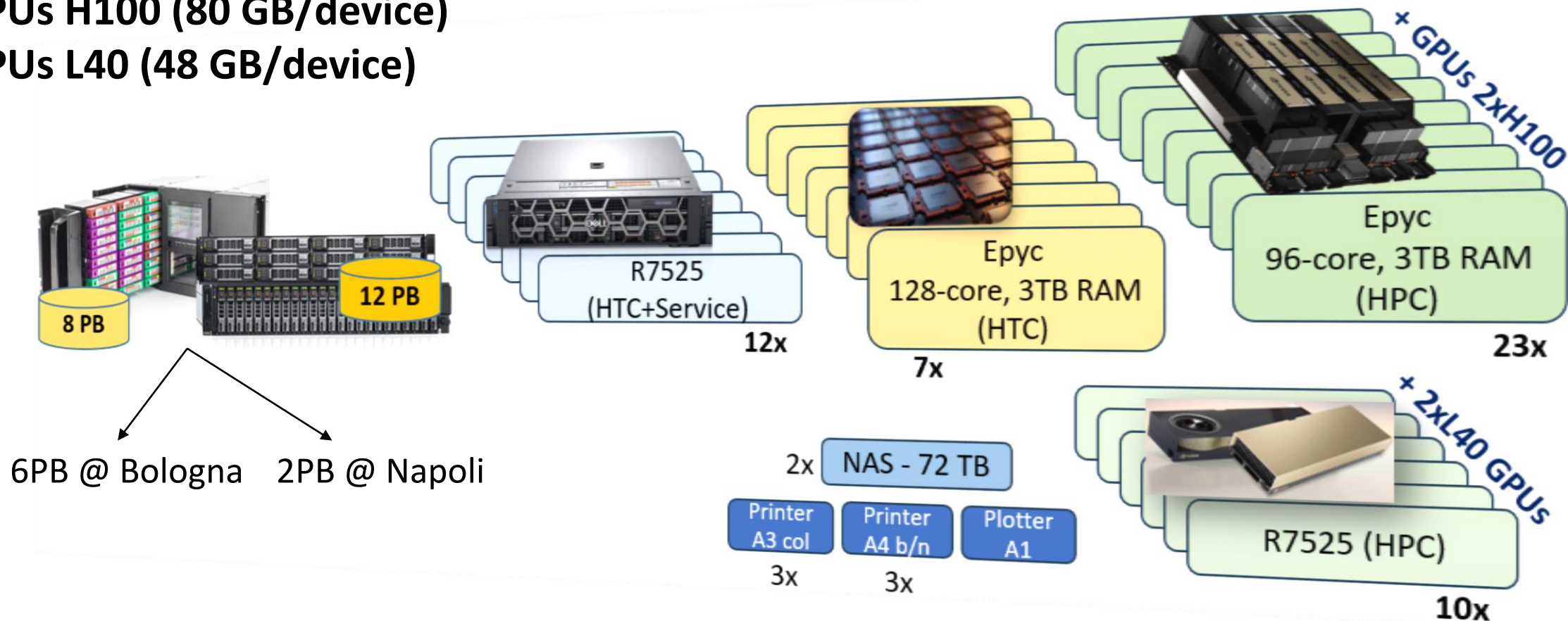
The Rubin-LSST@Italy in-kind Program is a long-term effort of INAF with evident benefits for the entire national community, in terms of scientific and technological exploitation of LSST data (same rights of US/Chilean community).

A light IDAC in-kind resource would be based on two concepts:

Data Access Center (partial mirroring of LSST data, mainly tabular data, but not only...)

Data Processing Center (hosting HPC/HTC scientific pipelines, exposing data products)

Potentially able to fulfil a heterogeneous science case ecosystem:

- **Observation data reduction and calibration;**
- **High resolution and wide field image analysis;**
- **Intensive and data-demanding machine/deep learning applications;**
- **Hybrid execution of multi-thread and many-core computing processes;**
- **Massive data cross-matching;**
- **On-demand data/applications Interoperability among user groups;**
- **Application developing/optimizing/testing;**
- **Hosting services for external communities (ELT, SKA, ALMA, VST, Euclid, LSST...);**

The *ADHOC* infrastructure is on the way and it represents a feasible and concrete pillar to support the ICT needs of the program...

...however, we hope that other complementary ICT solutions from INAF may prove equally willing to support the program in the medium and long term. *This could solve the ~14% of residual data rights deficit*

MESSAGES

Thank You!

These slides were
CHAT GPT FREE
generated by human intelligence