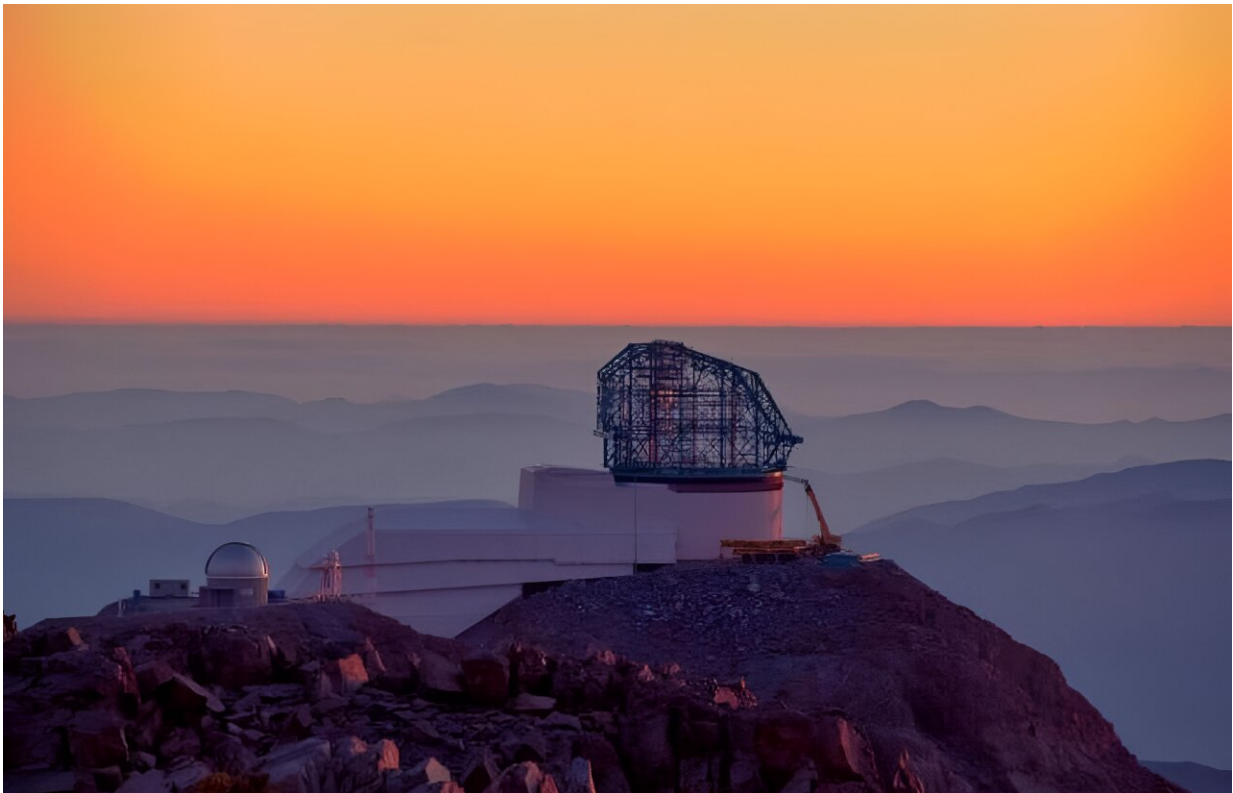


# Vera Rubin telescope will generate a mind-boggling amount of data, say astronomers

December 1 2023

---



The LSST, or Vera Rubin Survey Telescope, under construction at Cerro Pachon, Chile. Image Credit: LSST

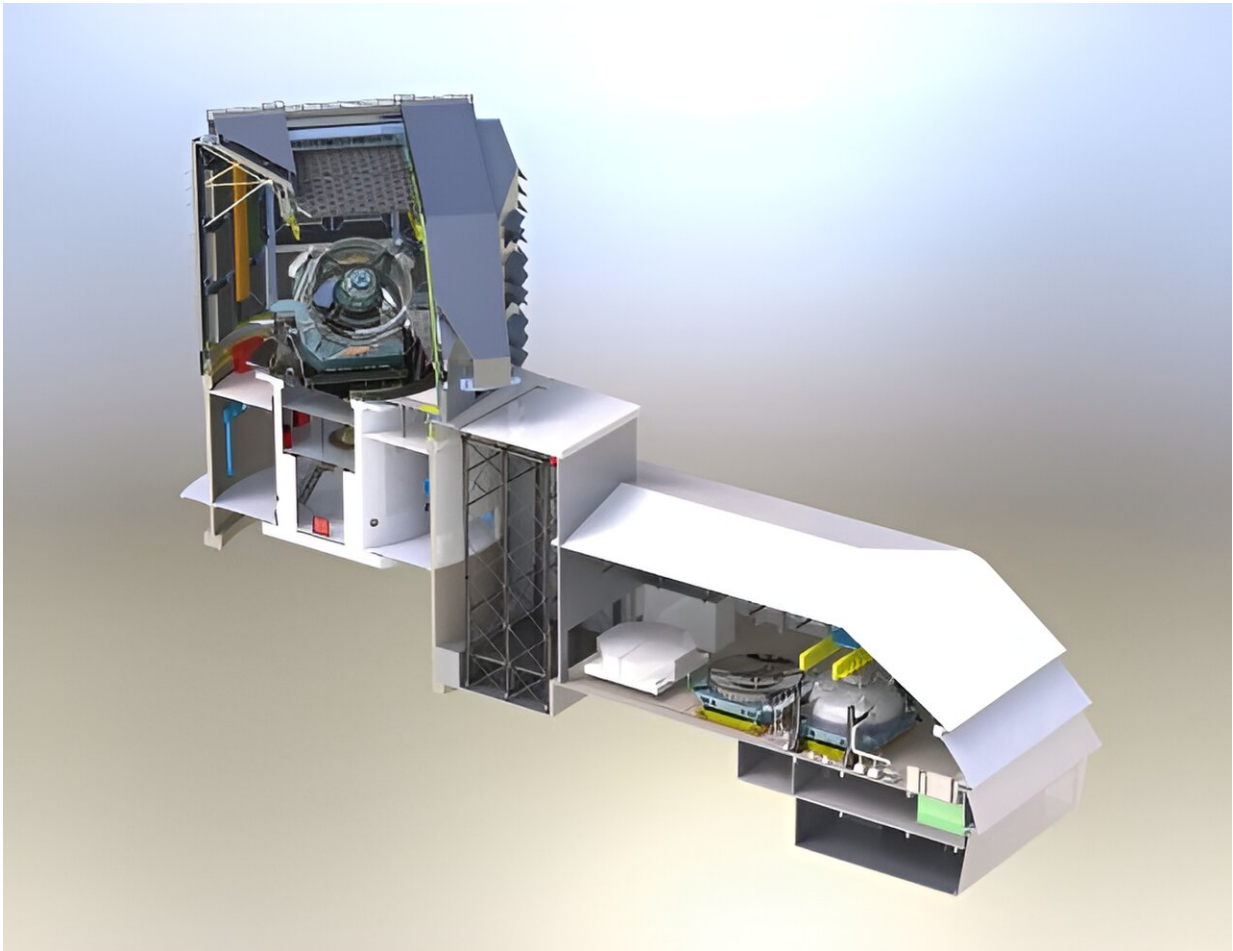
When the Vera C. Rubin Observatory comes online in 2025, it will be one of the most powerful tools available to astronomers, capturing huge portions of the sky every night with its 8.4-meter mirror and

3.2-gigapixel camera. Each image will be analyzed within 60 seconds, alerting astronomers to transient events like supernovae. An incredible 5 petabytes (5,000 terabytes) of new raw images will be recorded each year and made available for astronomers to study.

Not surprisingly, astronomers can't wait to get their hands on the high-resolution data. A new paper outlines how the huge amounts of data will be processed, organized, and disseminated. The entire process will require several facilities on three continents over the course of the projected 10-year-long survey.

The Rubin Observatory is a ground-based telescope located high in the Chilean Andes. The observatory's 8.4-meter Simonyi Survey Telescope will use the highest resolution digital camera in the world that also includes the world's largest fish-eye lens. The camera is roughly the size of a small car and weighs almost 2,800 kg (6,200 lbs). This survey telescope is fast-moving and will be able to scan the entire visible sky in the southern hemisphere every four nights.

"Automated detection and classification of celestial objects will be performed by sophisticated algorithms on high-resolution images to progressively produce an astronomical catalog eventually composed of 20 billion galaxies and 17 billion stars and their associated [physical properties](#)," write Fabio Hernandez, George Beckett, Peter Clark and several other astronomers in their [preprint paper](#) posted to *arXiv*.



Detailed cutaway render of the telescope model showing the inner workings.  
 Credit: LSST Project/J. Andrew

The main project for Rubin Observatory is the Legacy Survey of Space and Time (LSST) and researchers anticipate this project will gather data on more than 5 million asteroid belt objects, 300,000 Jupiter Trojans, 100,000 near-Earth objects, and more than 40,000 Kuiper belt objects. Since Rubin will be able to map the visible night sky every few days, many of these objects will be observed hundreds of times.

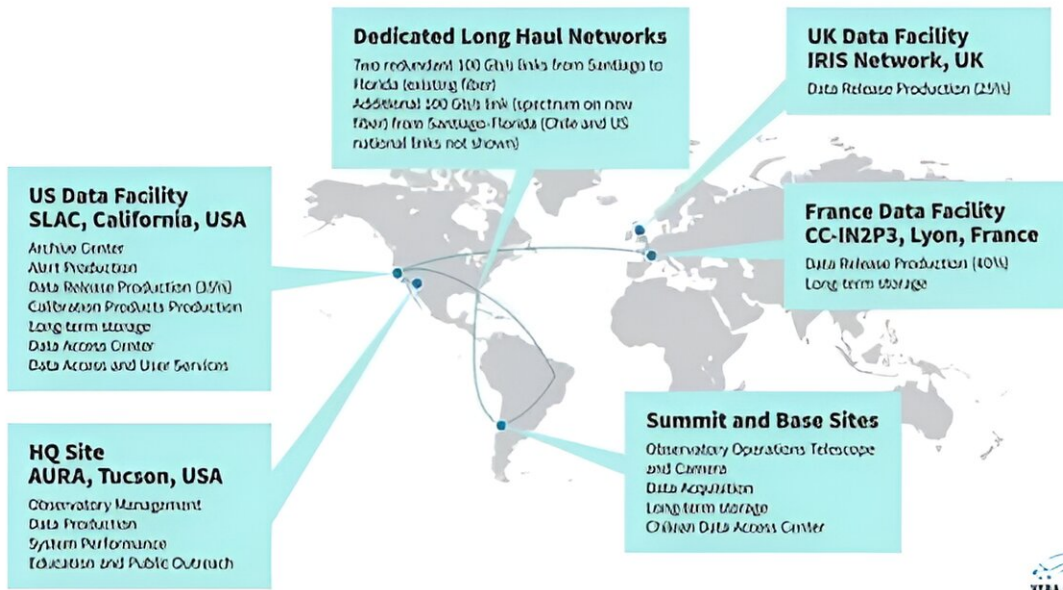
Because of the telescope's repeated observations, the enormous amount

of data will help calculate the positions and orbits of all these objects.

Images and data will immediately flow from the telescope to the Base Facility and Chilean Data Access Center in La Serena, Chile and then go to the three Rubin data facilities on dedicated high-speed networks connecting the sites: the French Data Facility CC-IN2P3 in Lyon, France, the UK Data Facility, IRIS network, in the United Kingdom and the US Data Facility and Data Access Center at SLAC National Accelerator Laboratory in California, U.S. There is also a Headquarters Site at the Association of Universities for Research in Astronomy (AURA) in Tucson, Arizona, U.S.

Once images are taken, they will be processed according to three different timescales: prompt, daily, and annually. The Hernandez et al paper outlines how raw images collected each observing night will be quickly processed (within 60 seconds), and objects that have changed brightness or position will generate and emit alerts for "transient detection."

For this process known as Prompt Processing, there will be no proprietary period associated with alerts, and they will be available to the public immediately, since the goal is to quickly transmit nearly everything about any given event, to enable quick classification and decision making. Scientists estimate Prompt Processing could generate millions of alerts per night.



Images flow from the Summit Site, where the telescope is located in Chile, to the Base Site and then to the three Rubin Data Facilities which collectively provide the computational capacity for processing the images taken by the Observatory for the duration of the survey. Credit: Vera Rubin Observatory

Daily products, released within 24 hours of observation, will include the images from that night. The annual campaigns will reprocess the entire image dataset collected since the beginning of the survey.

For each data release, there will be raw and calibration images in addition to science-ready images which have been processed with updated scientific algorithms. There will also be catalogs with the properties of all the astrophysical objects detected.

"The volume of released data products generated by the annual processing of the accumulated set of raw images is on average 2.3 times the size of the input dataset for that year and is estimated to reach more

than one hundred petabytes by the end of the survey," the astronomers wrote. They also said that over the 10 year-long survey the volume of data released for science analysis is estimated to increase by one order of magnitude.

The Rubin Observatory will utilize several kinds of data products and services for archiving and dissemination of the data to the various science collaborations. The paper says the Rubin LSST "Science Pipelines" are composed of about 80 different kinds of tasks, which are all implemented on top of a common algorithmic code base and specialized software. There is a feature called the Data Butler, which is the software system that abstracts the data access details (including data location, data format and access protocols).

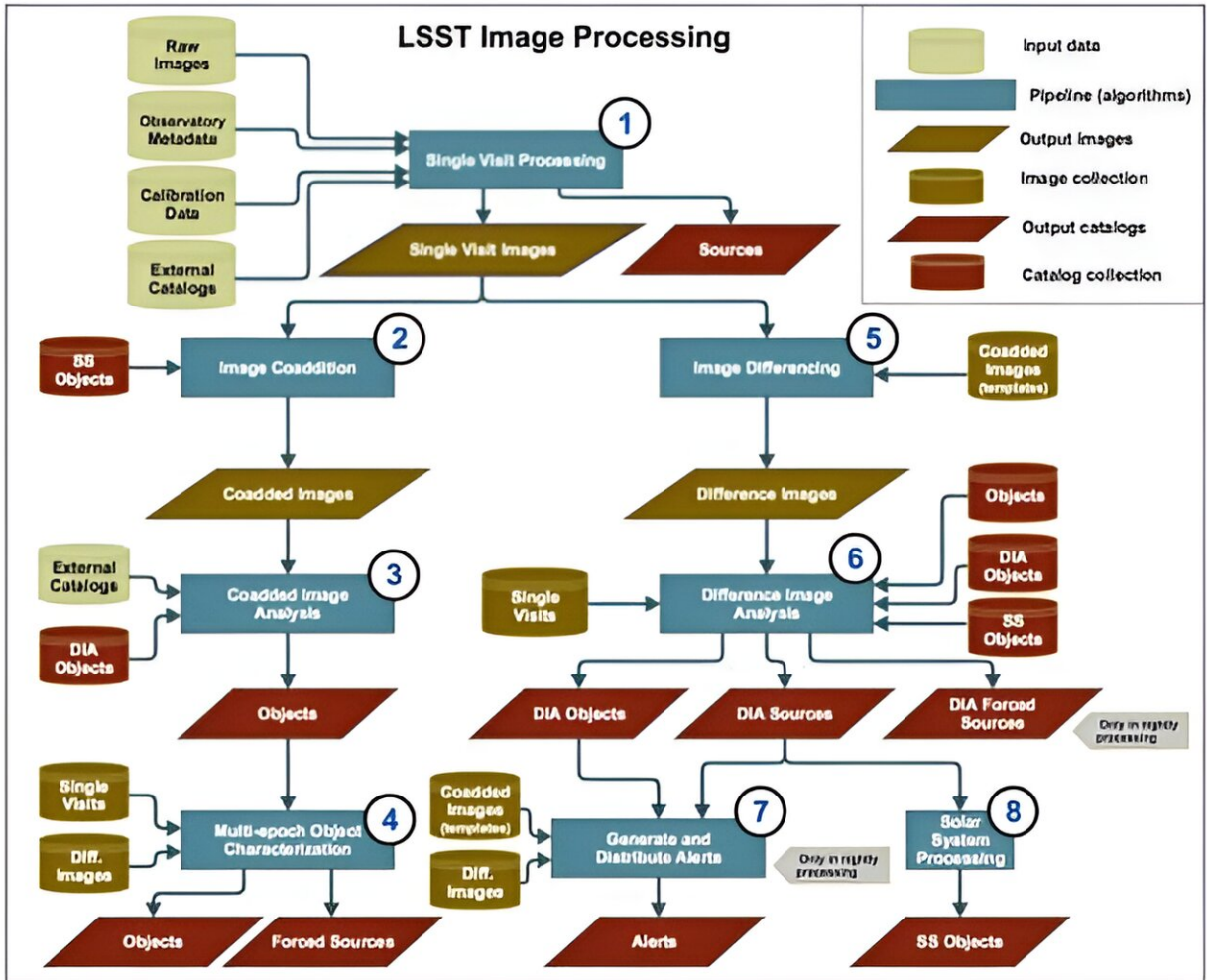


Illustration of the conceptual design of the LSST Science Pipelines for image processing. Credit: Hernandez et al.

Each year a data release will be produced and made available to science collaborations for use in studies in four main science pillars: probing [dark matter](#) and dark energy, taking inventory of solar system objects, exploring the transient optical sky and mapping the Milky Way.

This annual release will allow all the survey images taken to date to be reprocessed, combined and automatically measured to yield an

increasingly deep picture of the whole Southern sky, and a growing catalog of astronomical objects that captures how each one has changed over time. This annual data processing will be run at the three data facilities, with the final dataset assembled at SLAC and made available to astronomers and physicists via the Rubin Science Platform.

Right now, it is expected that Rubin Observatory data will become fully public after two years. The issue of how the public data can be accessed and how this access could be funded is still in the works.

**More information:** Fabio Hernandez et al, Overview of the distributed image processing infrastructure to produce the Legacy Survey of Space and Time, *arXiv* (2023). [DOI: 10.48550/arxiv.2311.13981](https://doi.org/10.48550/arxiv.2311.13981)

Provided by Universe Today

Citation: Vera Rubin telescope will generate a mind-boggling amount of data, say astronomers (2023, December 1) retrieved 3 May 2024 from <https://phys.org/news/2023-12-vera-rubin-telescope-generate-mind-boggling.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.