Adoption of AI at the Vera C. Rubin Observatory

AI Meets CI: Intelligent Infrastructure for Major & Midscale Facilities

Panel I: Adoption of AI in Major and Mid-scale Facilities (MFs)

CI Compass Virtual Workshop
2026-01-15

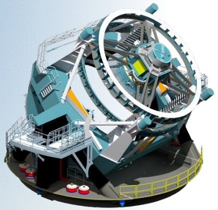
Leanne Guy

Vera C. Rubin Observatory
Data Management Scientist,
Associate Director, System Performance

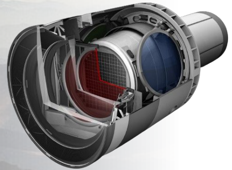
Vera C. Rubin Observatory

Start of operations in late 2025

Large-aperture, wide-field ground-based optical observatory at 2647 m on Cerro Pachón, Chile



Simonyi Survey Telescope: ∅ 8.4 m (effective 6.4m) main mirror | large aperture: f/1.234 | 350 ton | compact | settle of 4 sec fast slew & settle to adjacent field



Camera and Filters: 3.2 G pixels | 3.5° field of view, | 9.6 deg^2 | 6 filters, ugrizy, 320-1050 nm | 2 sec readout | 0.2”/pixel

Automated Data Processing System: 6.4 GB/ exposure | 20 TB/ night | | 10-yr dataset of 6 million images, 15 PB catalog | access via cloud-based science platform

Legacy Survey of Space and Time – LSST

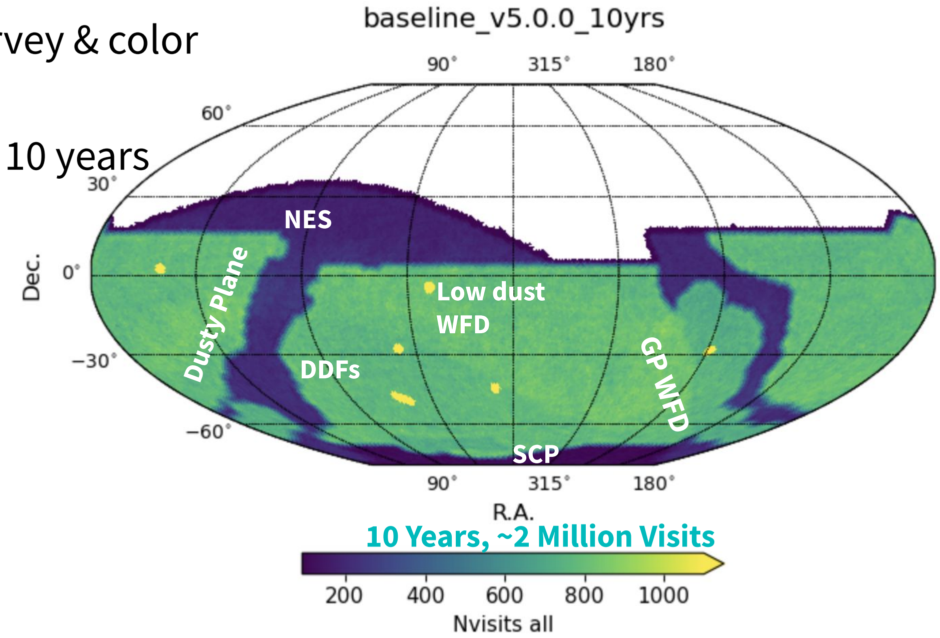
In the first 10 years of operation, the Vera C. Rubin Observatory will execute as its prime mission the Legacy Survey of Space and Time (LSST)

A “Wide-Fast-Deep” uniform optical/near-IR sky survey & color movie:

- Entire visible sky, (18K deg^2) every 3 nights for 10 years
- ~ 800 visits/per pointing,
- ugrizy bands (320 – 1050nm) , r~27.5 (36 nJy),

Vast astronomical dataset:

- 20 TB data, 10 million transient alerts night
- 60 GB raw images, 15 PB final catalog, 500 PB of image data products, 10 data releases.

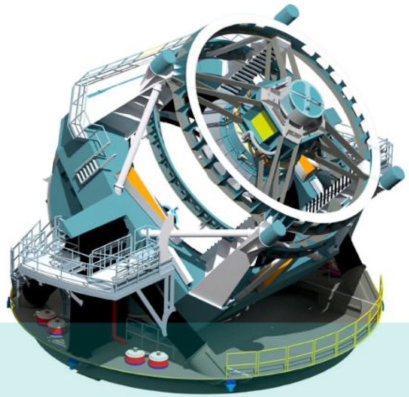


Expected Start: February – March 2026

Photons to Science

Raw Data: 20TB/night

Sequential 30s images covering the entire visible sky every few days



Prompt Data Products

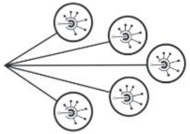
- Alerts incl. science, template and difference image cutouts — 60s via Alert Streams
- Catalogs of detections incl. difference images, transient, variable & solar system sources — 24h via Prompt Products
- Raw & processed visit images (PVIs), difference images — 80h via Image Services

Data Release Data Products

Final 10yr Data Release:

- Images: 5.5 million x 3.2 Gpixels
- Catalog: 15PB, 37 billion objects

via Data Releases



Community Brokers

Rubin Data Access Centres (DACs)

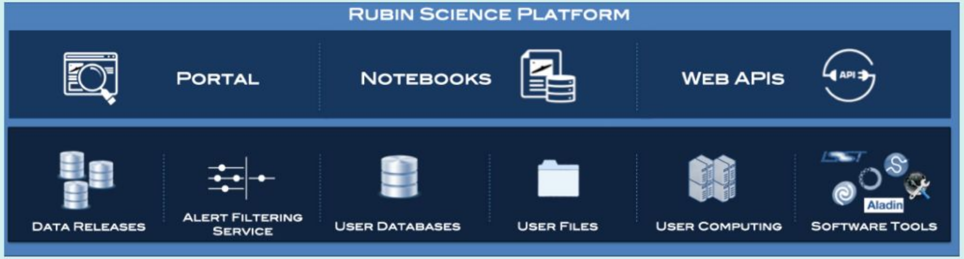
USA (USDF)
Chile (CLDF)
France (FRDF)
United Kingdom (UKDF)

Independent Data Access Centers (IDACs)

Access to proprietary data and the Science Platform require Rubin data rights

Rubin Science Platform

Provides access to LSST Data Products and services for all science users and project staff.



Credit: Leanne Guy

Rubin Key Performance Indicators

Delivered image quality: characterizes the spatial resolution; a measure of the information content of the images and includes contributions from static optics, windshake, telescope tracking, charge diffusion in sensors, dome seeing, and atmosphere. Impacts image depth, star-galaxy classification, astrometry, galaxy shapes, and blending.

Effective survey speed: the product of instantaneous etendue (ability to capture photons = effective area x field of view), observing efficiency, and system availability.

Normalized étendue: fE = fA * fS * fO * System Availability

- FOV Area (fA): Total live science pixel area under nominal conditions
- Sensitivity (fS): Visit rate including exposures, slews, readout, and filter changes
- Observing Efficiency (fO): Losses from weather and system downtime

AI for Wave-front Estimation with the Rubin AOS

Rubin uses an **Active Optics System(AOS)** to correct the alignment and mirror surface perturbations introduced by gravity and temperature gradients in the optical system to ensure precision-quality images.

Rubin has developed a **deep-learning model for wavefront estimation in the AOS pipeline,** which is faster and more robust than the baseline algorithm.

Using AI for Wave-front Estimation with the Rubin Observatory Active Optics System

John Franklin Crenshaw, Andrew J. Connolly, Joshua E. Meyers, J. Bryce Kalmbach, Guillem Megias Homar, Tiago Ribeiro, Krzysztof Suberlak, Sandrine Thomas, and Te-Wei Tsai

Published 2024 January 31 • © 2024. The Author(s). Published by the American Astronomical Society.

The Astronomical Journal, Volume 167, Number 2

Citation John Franklin Crenshaw *et al* 2024 *AJ* **167** 86

DOI 10.3847/1538-3881/ad1661

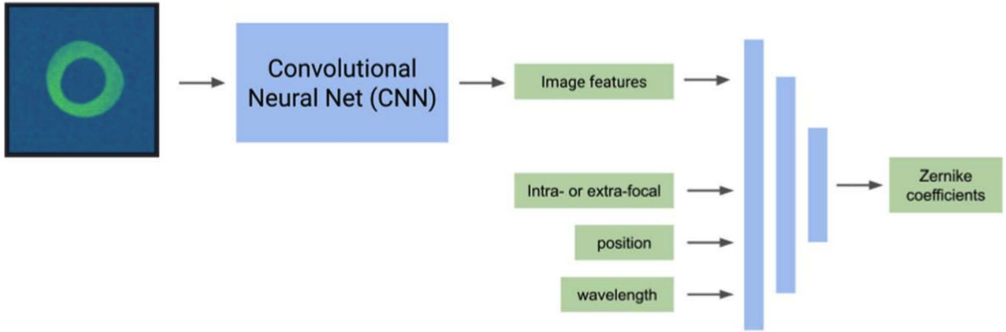
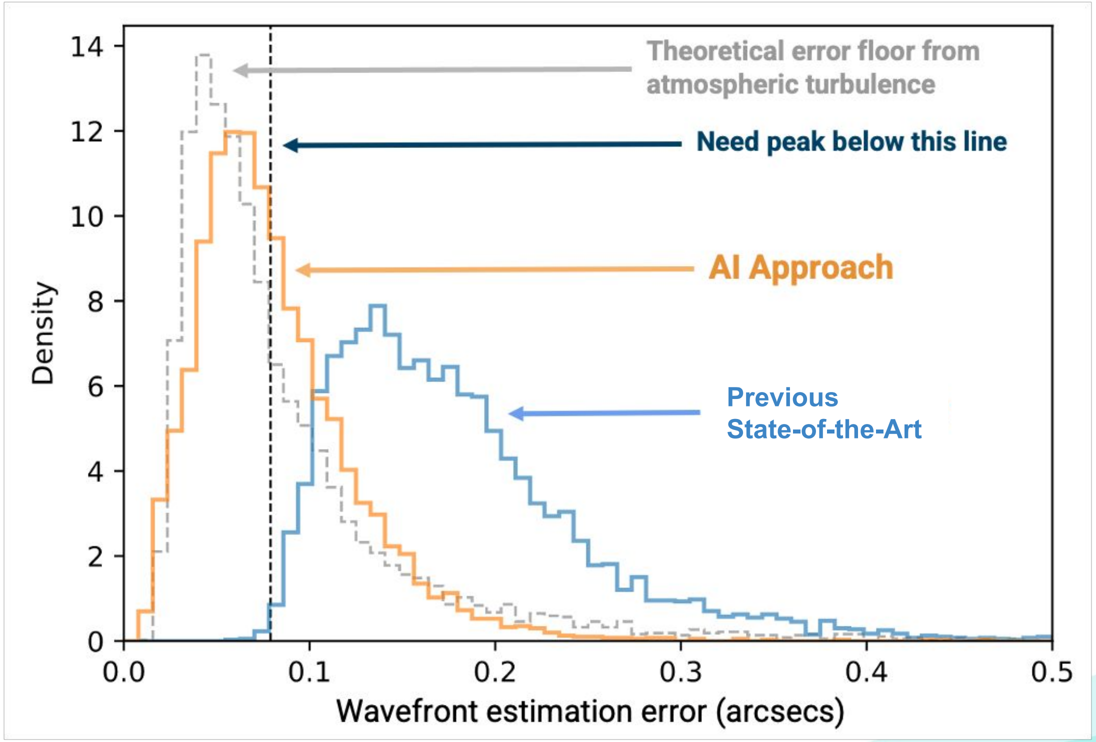


Figure 2. Schematic of the network architecture. A single donut is passed through a convolutional neural network (CNN), generating a set of image features. These features are concatenated with a flag that indicates if the donut is intra-/extra-focal, the field angle of the source, and the effective wavelength of the observation. A dense neural network then estimates the Zernike coefficients associated with the optical wave front.

AI for Wave-front Estimation with the Rubin AOS

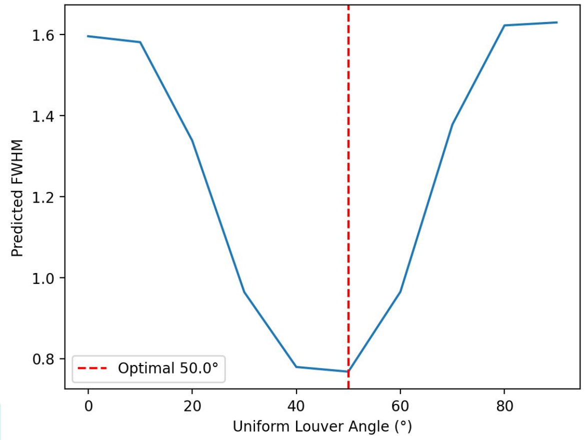
The Rubin Deep Learning approach achieves the atmospheric error floor even in challenging conditions, with up to 40× speed improvement and significantly lower errors under vignetting and blending.



Louver Optimisation using Machine Learning

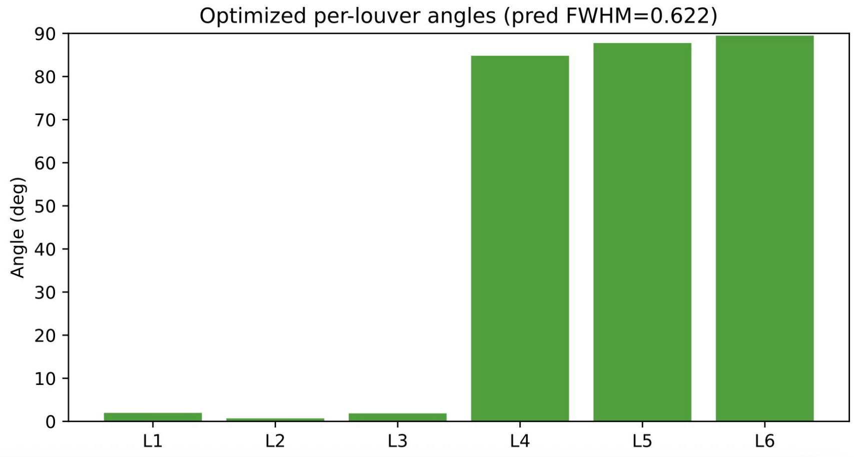
Louvers control airflow and temperature to reduce dome turbulence and optimize delivered image quality (FWHM). Marina Pavlovic is developing ML models to predict delivered image quality (FWHM) from environmental conditions and suggest optimal louver angles, with initial tests using LightGBM, random forests, and feedforward neural networks (MLPs).

Light GBM



Uniform louver angle 50.0º (predicted FWHM 0.768)

FFNN



Optimized per-louver angles (predicted FWHM 0.622)

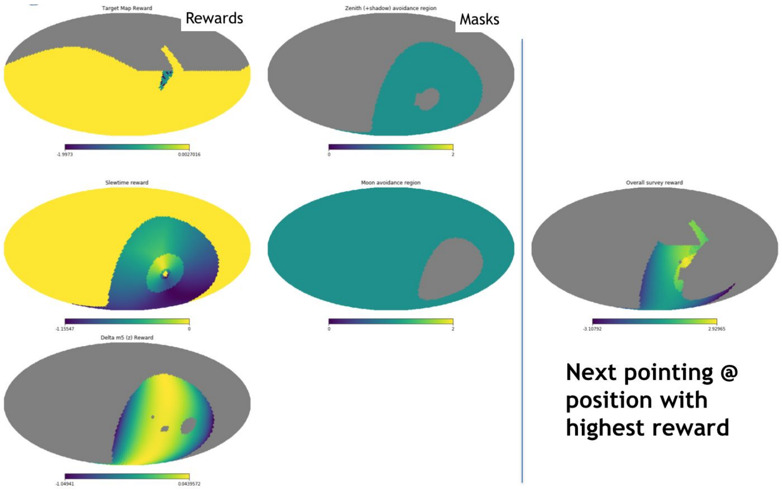
AI in the Rubin Feature-Based Scheduler (FBS)

Real-time scheduling: Builds on the baseline survey design to optimize the sequence of nightly observations in real time in response to evolving environmental conditions (e.g., weather) and system state, while enforcing and balancing multiple scientific priority constraints.

Decision & learning framework: Uses a Markov Decision Process (MDP) to rank candidate observations, with features tracking state, basis functions scoring options, and surveys combining scores with priority tiers. While MDPs are not AI, basis-function weights can be AI-learned. competing science metrics.

A Framework for Telescope Schedulers: With Applications to the Large Synoptic Survey Telescope

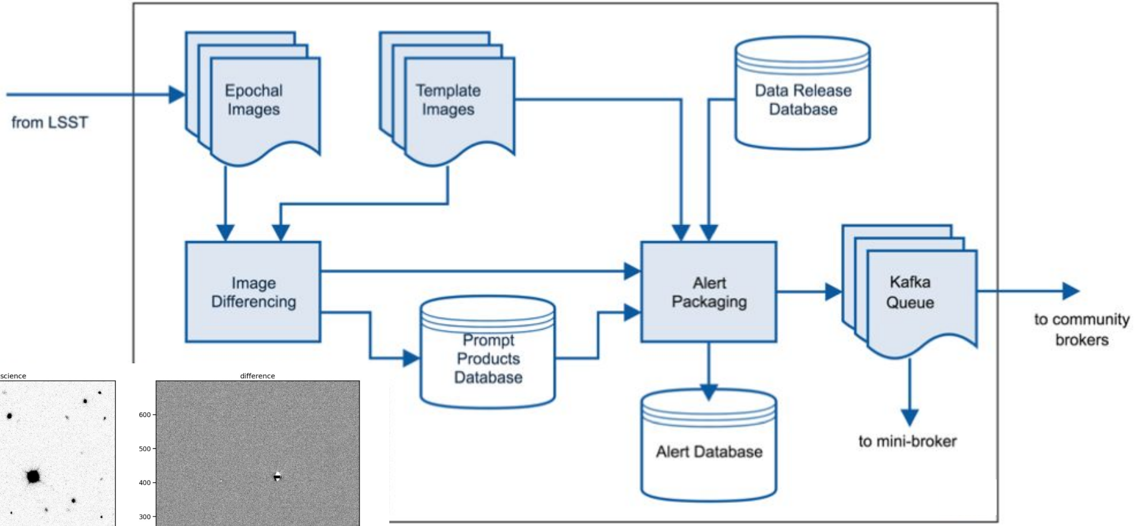
Elahesadat Naghib[1], Peter Yoachim[2], Robert J. Vanderbei[1], Andrew J. Connolly[2], and R. Lynne Jones[2]
[1] Department of Operations Research and Financial Engineering, Princeton University, Princeton, NJ 08540, USA; enaghib@princeton.edu, rvdb@princeton.edu
[2] Department of Astronomy, University of Washington, Box 351580, U.W., Seattle, WA 98195, USA; yoachim@uw.edu, ajc26@uw.edu, ljones@astro.washington.edu
Received 2018 October 5; revised 2018 December 11; accepted 2019 January 7; published 2019 March 22

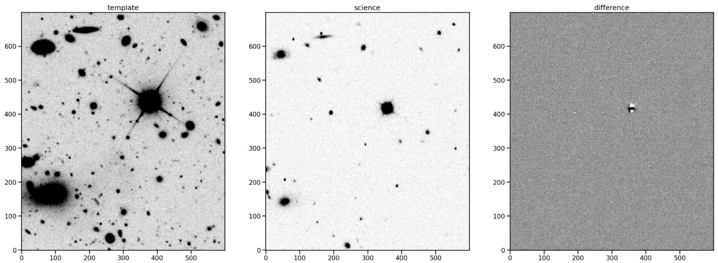


FBS selects fields by maximizing a reward balancing open-shutter time and slew losses.

Rubin Alert Production

Rubin Alert Production System: Processes nightly LSST images in near real-time to detect transients and variable sources, generating rapid alerts with measurements and context for community follow-up.

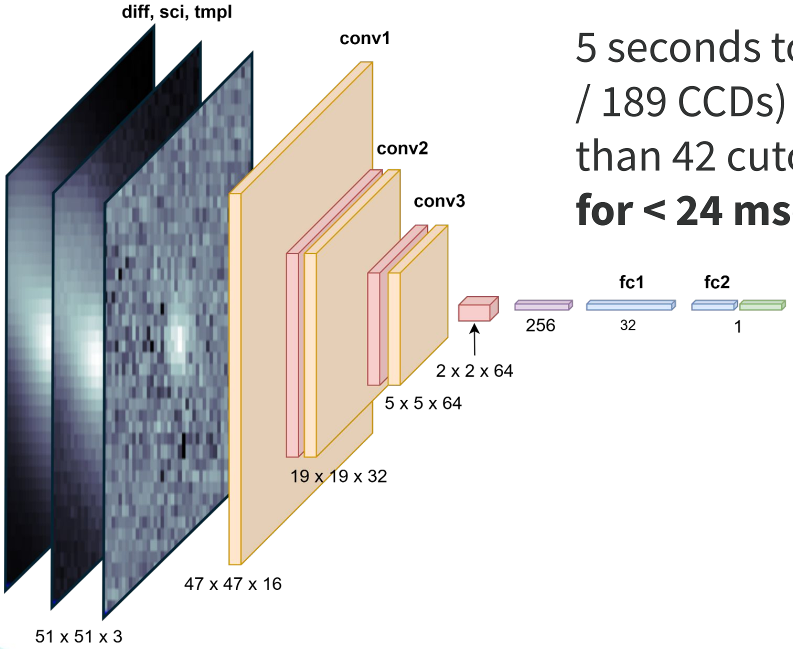


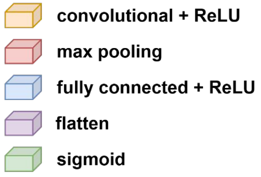


dataId = {'instrument': 'LSSTComCam', 'detector': 6, 'visit': 2024111500287}

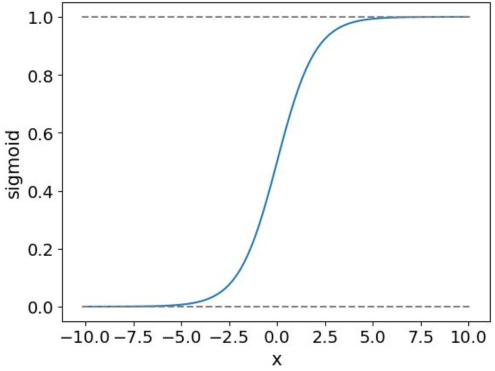
Distinguishing real astrophysical transients from spurious detections in automated survey data is challenging and we are investigating AI methods to assist

Fast CNN to help eliminate false positives



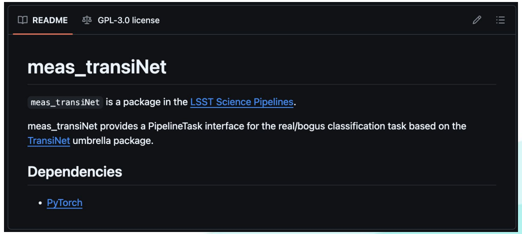


5 seconds to process (40,000 stamps / 189 CCDs) ~ 212 stamps ⇒ faster than 42 cutouts/second ⇒ **design for < 24 msec per cutout**



Output gives the probability of being real – **the reliability score.**

This ML reliability model was developed with the aim to meet the latency requirements for Rubin Alert Production on CPUs


README GPL-3.0 license

meas_transiNet

`meas_transiNet` is a package in the LSST Science Pipelines.

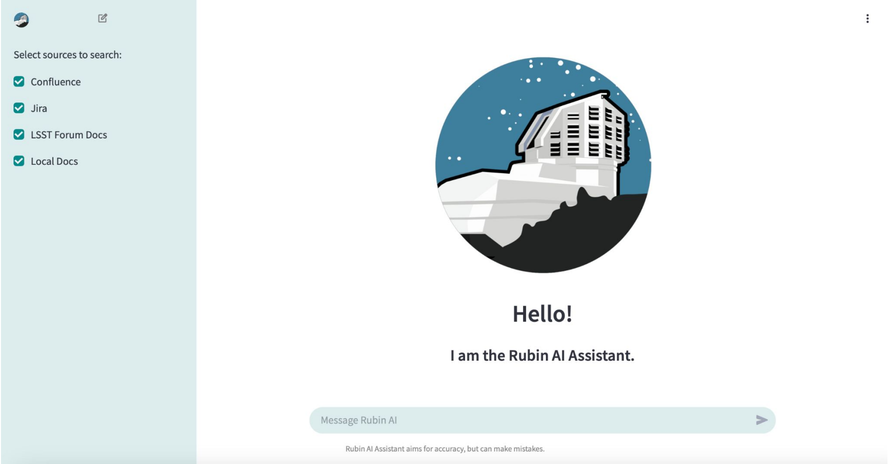
meas_transiNet provides a PipelineTask interface for the real/bogus classification task based on the TransiNet umbrella package.

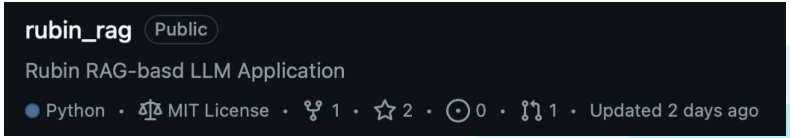
Dependencies

- PyTorch

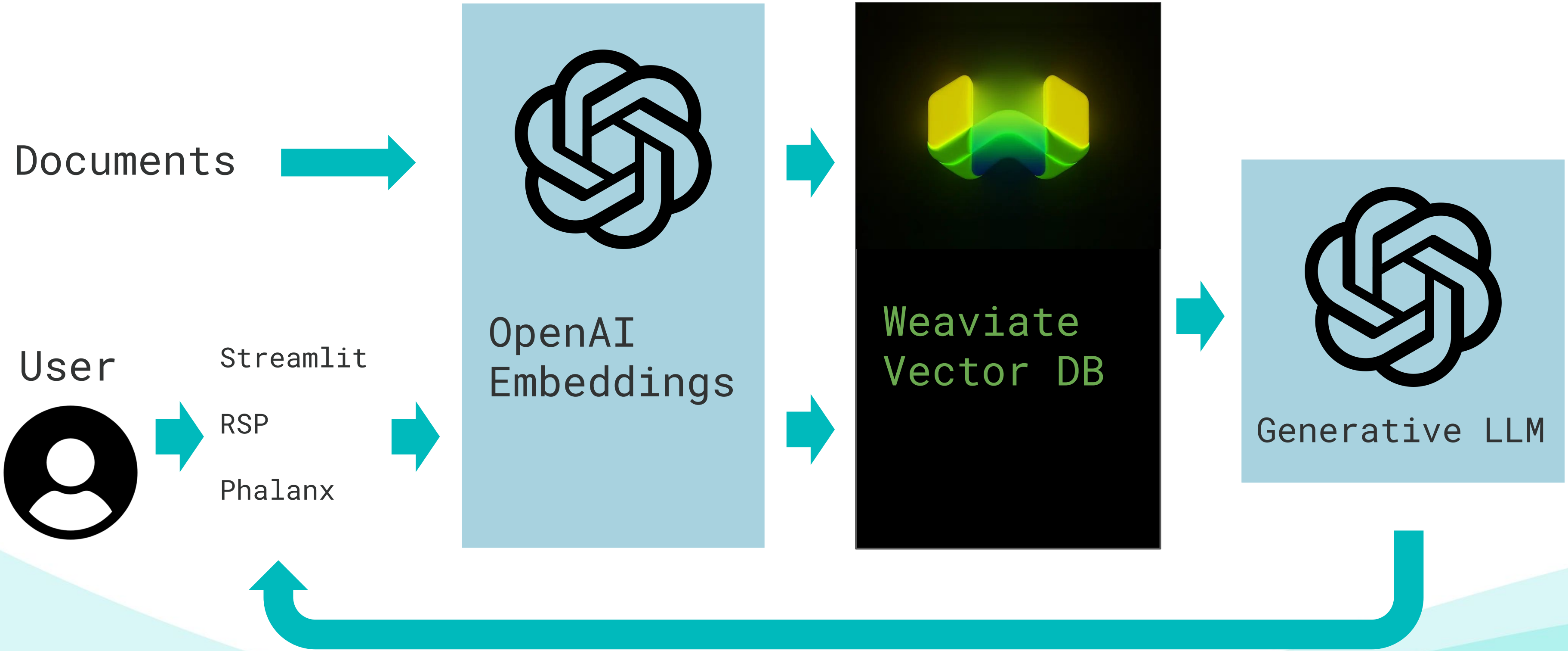
Rubin Generative AI ChatBot for Community Support

- Rubin is developing a prototype Retrieval-Augmented Generation (RAG) chatbot to support observatory staff, collaborators, and the community.
- It answers Rubin-specific questions (e.g., Butler code, pipelines, data products, Jira) that standard LLMs struggle with.
- Using RAG, the chatbot augments existing LLMs with domain knowledge,
- The current stack includes LangChain, Weaviate, OpenAI, Streamlit,
  And uses Rubin's Phalanx deployment Model
- Prototype deployed on Rubin google cloud data access platform





Design of a Rubin_RAG: Architecture & Components



Laying the Foundation for AI-Assisted Discovery

Where we are: AI adoption today at Rubin is largely about scale — helping us manage volume, velocity, and complexity, and making data-intensive workflows faster and more robust. This acceleration is a critical foundation for working effectively with big scientific datasets.

Where we're going: As methods improve, AI at Rubin will move beyond speed and robustness to enable the discovery of rare phenomena and anomalies — revealing new science that was not previously possible. AI accelerates discovery and expands what we can learn from data, while amplifying, not replacing, domain expertise by combining human understanding with machine scale.

This is just a snapshot of some of the projects using AI at Rubin Observatory.

There are many more in various stages of development, e.g. predictive maintenance, photometric redshift estimation, interpolation of PSFs between stars, Scarlet2, Gaussian Processes fit for WCS to better account for atmospheric turbulence, Scarlet for deblending, etc.

Many staff are using AI coding tools integrated into their workflow, e.g. github copilot, ChatGPT.