VERA C. RUBIN OBSERVATORY

Data Management All Hands Meeting 2023

William O'Mullane Data Management
Project Manager / Associate Director
AURA/Rubin Observatory

13th March 2023

NSF NOIR Lab AURA SLAC U.S. DEPARTMENT OF ENERGY

Meeting Conduct

- Remember we have a code of conduct for meetings Document-28973 and one for communications Document-24920.
- Shared responsibility for health of the community and thus meeting atmosphere Respect - consider a variety of viewpoints and approaches:
  - notbuttinginconstantlywhenotherpeoplearespeakingbecausewedisagree Honesty - be prepared for some civil discussion
- To respect every ones time a good meeting should have:
  - An agenda which should be largely adhered to
  - Recorded Minutes with Decisions and Actions (which may be confluence Jira)
- In this online meeting we will have a moderator in each session
  - Use raise hand feature of zoom for questions (as moderator requests) - **Tim Jenness**
  - Questions may be best in DM slack channel with threads (rather than zoom chat)

Agenda

1. DM/DP highlights and plans (m)
2. Summit architecture - KT (15-20m)
3. Lightning talks (15-30m)
4. 15:55 Photo on left patio (my left)! And coffee break
5. 16:30 Phalanx - Frossie (15m)
6. Sasquatch - Angelo (15m)
7. Any other questions
8. 17:15 Enjoy La Serena

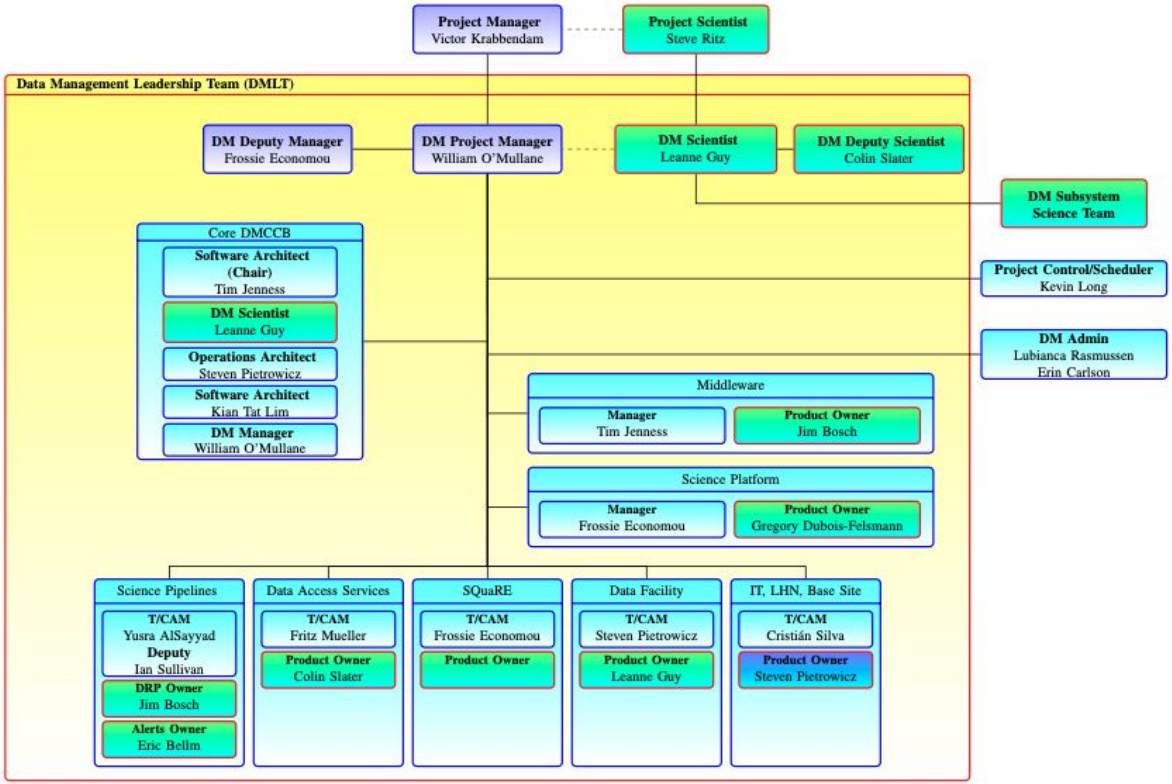
Data Management
Organization

Management team remains stable.

Lubi (admin) added

Large gap in Square Product ownership.

LDM-294 updated but always behind

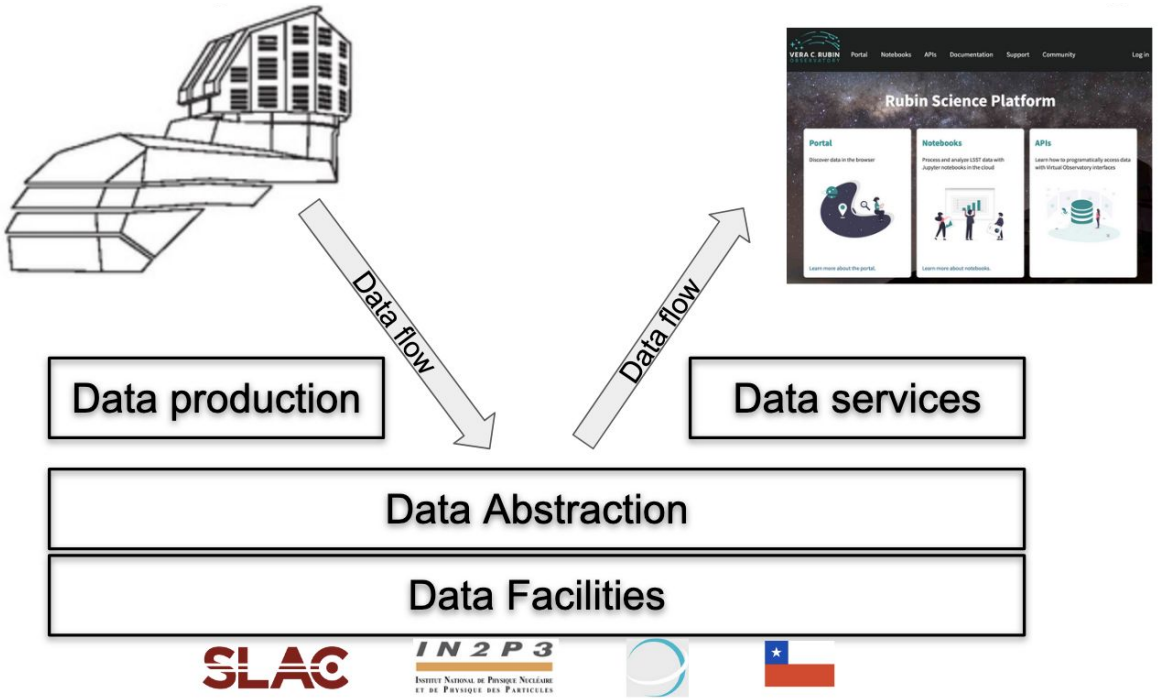


Data Management OPS organization

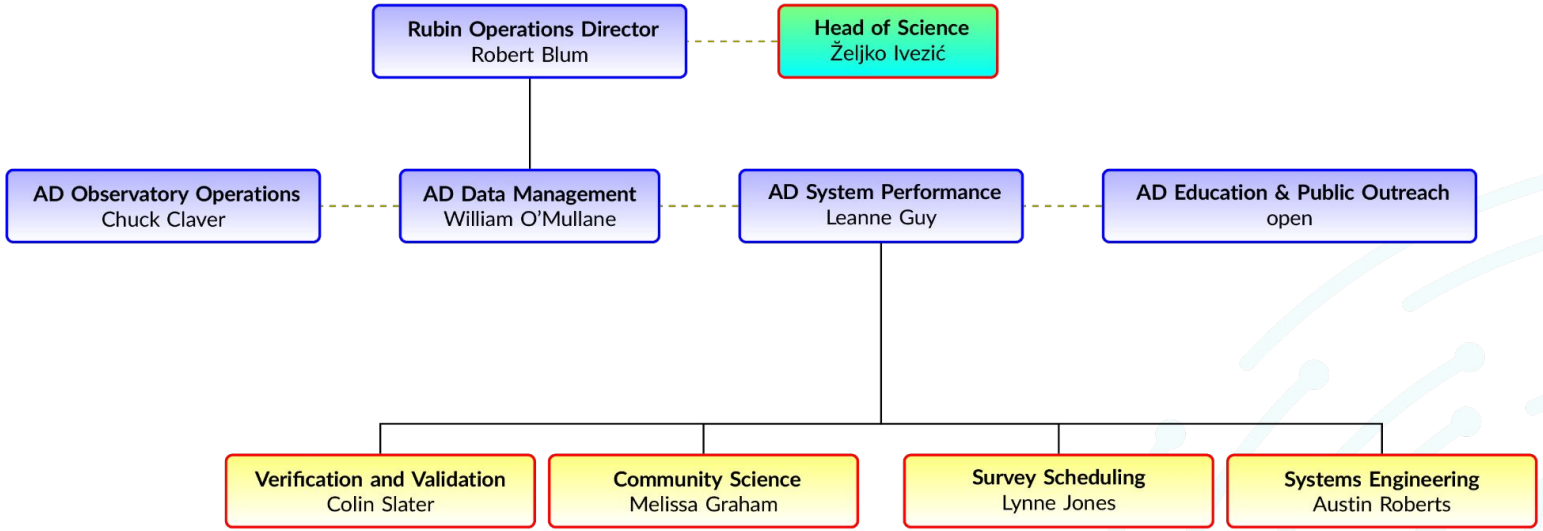
See RTN-046

Five areas of oversight aligned with core activities

- Data Production
- Data Services
- Data Abstraction
- Data Facilities
- Data Acquisition



System Performance Organisation



Rubin Observatory System Performance department is responsible for ensuring that the LSST as a whole is proceeding with the efficiency and fidelity needed to achieve its science requirements at the end of the 10-year survey

Welcome to new DM people…

- Bruno Sanchez (AP)
- Kai Koehler (SQuARE)
- Omari Paul (Data Curation)
- Ramya Eranna (Infrastructure)
- Peter Vaucher (Infrastructure)
- Sreevani Jarugula (System Performance and Infrastructure)
- Becky Nevin (System Performance and Infrastructure)

We also had Staff Highlights

Carlos Barría

DevOps IT Chile | Joined June 2022

Along with the IT Chile team, I help prepare and maintain in excellent condition, the servers and clusters that work and process Scientific Data.

I am proud of my work in automating the provisioning, configurations, updates and services of the large number of servers that we manage, whether they are virtual, bare metal or Kubernetes.

In my free time I play the guitar, play online video games and enjoy cooking.

Junto con el
emos en exc
ters que trab

Automatizar
ciones, actu
de servidore
metal o Kub

En mi tiempo
online y disfr

Nima Sedaghat

Deep-Vision Scientist, AI Lead at LSST AP | Científico de Visión Profunda, Principal de IA en LSST AI
Joined November 2021

I help lead a couple of projects working on development of deep learning/AI-based methods for the LSST Science Pipelines. My current focus is on transient hunting, or more exotically put, supernova hunting.

While I was still doing my PhD in Computer Vision a few years ago, I thought of giving my childhood passion (astronomy) a try. I applied computer vision techniques on astronomical data, developed an AI-based algorithm for supernova hunting, and published it under the name TransiNet. Now, a few years later, I have joined the great team of Rubin scientists, and am implementing TransiNet as a key part of the cutting edge observatory.

As a teenager I was obsessed with astronomy, to the extent that I made a 15 cm telescope mirror (by grinding and polishing) in the basement of our house. But I was also interested in hiking and climbing, so much that I fell off a 17 m high cliff at the age of 18. Anyway, I recovered and continued studying Electrical Engineering at university—notice anything strange? Now I am working at the cross-section of signal processing, deep learning, and astronomy, and am grateful for whatever caused this not-so-straight path!

Ayudo a liderar un par de proyectos que trat
rollo de métodos basados en aprendizaje pr
los Dato ductos Científicos de LSST. Mi enfoc
búsqueda de transitorias, o dicho de manera más exótica, la
búsqueda de supernovas.

Cuando estaba haciendo mi doctorado en Computer Vision
hace unos años, pensé en darle una oportunidad a mi pasión
de la infancia (la astronomía). Apliqué técnicas de computer
vision en datos astronómicos, desarrollé un algoritmo
basado en IA para la bú
con el nombre de Trans
he unido al gran equipo
mentando TransiNet co
rio de vanguardia.

Cuando era adolescent
tronomía, hasta el punt
(pulido) de 15 cm en el
interesaba el senderism
un acantilado de 17 m
modos, me recuperé y
en la universidad, ¿nota
do en la sección transve
aprendizaje profundo y
lo que haya causado es

Fritz Mueller

Research Database Technical Lead | Jefe Técnico de la Base de Datos de Investigación
Joined December 2014

I lead a small development team at SLAC working on design and implementation of high-performance database systems that are necessary to keep up with Rubin's ambitious scale and pace. These systems include Qserv (a scalable SQL database for released survey catalogs) and the Alert Production Database (time-critical online support for Rubin's alert production pipelines). Lately I have also been helping out with the planning and coordination of Rubin's US Data Facility at SLAC.

During my tenure at SLAC, we've seen the Qserv system mature from an early prototype to a robust and reliable production system. Recently, we substantially re-factored Qserv's build, configuration, and operating environments to enable its modern life as a Kubernetes-based distributed system. Instances are now running in the Google cloud and at CC-IN2P3, serving data previews to hundreds of science users as a back end component of the Rubin Science Platform.

I very much enjoy troubleshooting and repair of anything, from clocks and radios to vintage guitar amplifiers, arcade games, keyboards, test equipment, and minicomputers (many of these repair adventures are technically documented on a blog I keep at https://fritzm.github.io). I am an avid retro-computing enthusiast, and maintain an operational PDP-11/45 in my basement. I also have a background in professional audio, and in a previous professional life worked designing and building mixing consoles for music, stage, and film production, on embedded systems for digital audio workstations, and on firmware for sampling synthesizers. I have played electric and upright bass in a few bands around the San Francisco Bay Area.

Estoy a cargo de un pequeño equipo de desarrollo en SLAC que traba
ja en el diseño y la implementación de sistemas de bases de datos de
alto rendimiento los cuales son necesarios para seguir el ambicioso
ritmo y escala de Rubin. Estos sistemas incluyen Qserv (una base de
datos SQL escalable para catálogos de investigaciones emitidos) y la
Base de Datos de Producción de Alertas (soporte en línea crítico para
los datoductos de producción de alertas de Rubin). Últimamente, tam
bién he estado ayudando con la planificación y coordinación de la in
stalación de Datos en Estados Unidos de Rubin en SLAC.

Durante mi permanencia en SLAC, he visto como el sistema Qserv
maduró desde un prototipo inicial a un sistema de producción robusto
y confiable. Hace poco realizamos una sustancial refactorización de
los entornos operativos, de configuración y de construcción de Qserv
para permitir la modernización a un sistema distribuido basado en Ku
bernetes. Las instancias están siendo ejecutadas en la nube de Google
y en CC-IN2P3, brindando vistas previas de datos a cientos de usuarios
científicos como un componente back-end de la Plataforma Científica
de Rubin.

Disfruto resolver problemas y reparar cosas, desde relojes y radios
hasta amplificadores de guitarra antiguos, juegos de arcade, teclados,
equipos de prueba y minicomputadoras (varias de estas aventuras de
reparación están técnicamente documentadas en un blog https://-
fritzm.github.io). Soy un ávido entusiasta de la informática retro y
mantengo un PDP-11/45 operativo en mi sótano. También tengo expe
riencia en audio profesional, y en mi vida laboral anterior, trabajé
diseñando y construyendo mesas de sonidos para la producción de
música, teatro y películas, en sistemas embebidos para estaciones de
audio digital y en firmware para sintetizadores sampler. Además, he
tocado el bajo y el contrabajo en algunas bandas en el Área de la
Bahía de San Francisco.

Hernán Stockebrand

IT Network Engineer | Ingeniero en Redes
Joined March 2021

I work as a Network Engineer, where I design, maintain and continuously make improvements, interconnecting all the observatory systems, from the User Network control systems, to the future traffic of data that will be produced by the 3.2 gigapixel camera.

Network Automation: Automation of maintenance tasks, configuration and updating of our network infrastructure.

Awarded astrophotographer with publications in newspapers, magazines and television interviews thanks to my images and knowledge of observational astronomy. I like to automate things; in my home, all lighting, heating, TVs are automated, which can be controlled by voice.

Trabajo como Inger
mos, mantenemos
nonectado todo lo
la Red de Usuarios,
trafico de los datos
camara de 3.2 gigap

Network Automatio
mantencion, config
infraestructura de n

Astrofotografo, con
y revistas, entrevist
genes y conocimien
al. Me gusta automa
tomatizadas todas
que pueden ser con

Fred Moolekamp

Research Scientist | Investigador Científico
Joined August 2016

I'm part of the team that works on the science pipelines algorithms to calibrate and analyze images taken by the Rubin Observatory. Over the 10 year survey each patch of the sky will be revisited ~2000 times, and combining those images allows us to see fainter (and farther) things. This results in images that become increasingly crowded, with galaxies and stars overlapping one another and making it difficult to make measurements on any one object. So my primary job is to work on software to "deblend" those stars and galaxies so that we can attempt to make measurements on them as if they were isolated in the sky.

Together with Peter Melchior and the assistance of other postdocs and faculty members, I co-developed the "scarlet" software package to deblend stars and galaxies. The package built on the successful deblending algorithm used in the Sloan Digital Sky Survey, expanded it by working with images taken at different bands (wavelengths), and implemented an improved optimization algorithm. In 2021 this became the primary deblender used in the science pipelines and it will continue to be modified and improved over the next few years.

My main interest is in physics and math education, so I enjoy giving talks, lectures, and developing apps to aid in communicating science at all levels. When I'm not thinking about science I'm probably either exercising, trying to enjoy the outdoors with my family (as weather in western NY permits), or starting various art projects that will never actually be finished.

Formo parte de
ductos científic
por el Observat
cada parte del
al combinar es
(y más lejanas)
cada vez más c
que dificulta re
trabajo princip
esas estrellas y
ciones de ellas

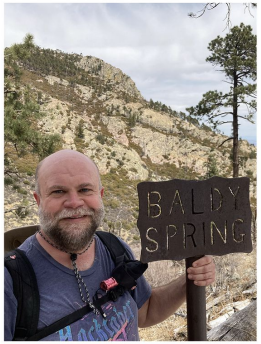
Junto con Pete
miembros de la tacuatao, desarrollamos en conjunto el paquete de software "escarlata" para separar estrellas y galaxias. El paquete se basó en el exitoso algoritmo de eliminación de mezclas utilizado en Sloan Digital Sky Survey, el cual fue ampliado al trabajar con imágenes tomadas en diferentes bandas (longitudes de onda) y se implementó un algoritmo de optimización mejorado. En 2021, este se convirtió en el principal separador de señales que se superponen utilizado en los datoductos científicos y seguirá siendo modificado y mejorado en los próximos años.

Mi principal interés es la educación en la física y matemáticas, por lo que disfruto dando charlas, conferencias y desarrollando aplicaciones para ayudar a comunicar la ciencia a todo nivel. Cuando no estoy pensando en la ciencia, probablemente estoy haciendo ejercicio, tratando de disfrutar del aire libre con mi familia (siempre que el clima en el oeste de Nueva York lo permita) o comenzando varios proyectos de arte que en realidad nunca se terminarán.

Thanks for participating in Staff Highlights

Adam Thornton

Devops Engineer | Ingeniero Devops
Joined November 2016



I work for the Science Quality and Reliability Engineering team (SQuaRE) within Data Management. I have mostly been working on the Notebook Aspect of the Rubin Science Platform (RSP), but am also now picking up some telemetry for Science Platform components. The Notebook Aspect is designed to provide a pleasant interface for quick tests of hypotheses across smallish, but arbitrary, slices of data. I've also been working on standardizing the construction and configuration of Science Platform components, to make it easier to extend the RSP with new services and applications.

If you use RSP notebooks, you're using the stuff I've been working on. At the end of 2022 and the beginning 2023, we wanted to make spawning user labs under heavy load more reliable. After quite a lot of effort, I was able to address everyone's concerns (it was a major architectural change) and get the asynchronous Kubespawner accepted upstream as the standard implementation--so now Jupyter maintains it, and the support burden is not solely on us.

I restore and play with old computers and video game systems. My oldest working console is an Atari Super Pong from 1976; I have quite an extensive collection of 8, 16, 32, and 64-bit home computers and consoles from the 1970s through today. I have three large dogs on whom I dote. I make and enthusiastically consume craft cocktails, and I have recently taken up motorcycle riding, though not at the same time.

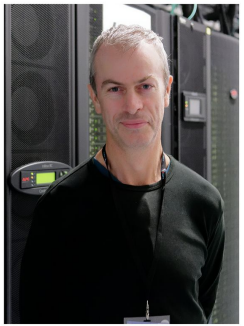
Trabajo para el equipo de Ingeniería de Confiabilidad y Calidad Científica (SQuaRE por sus siglas en inglés) dentro de Gestión de Datos. He estado trabajando principalmente en el Notebook Aspect de la Plataforma Científica de Rubin (RSP), pero ahora también estoy adquiriendo algo de telemetría para los componentes de la Plataforma Científica. El Notebook Aspect está diseñado para proporcionar una interfaz agradable para pruebas rápidas de hipótesis en porciones de datos pequeñas, pero arbitrarias. También he estado trabajando en la estandarización de la construcción y configuración de los componentes de la Plataforma Científica, para facilitar la extensión del RSP con los nuevos servicios y aplicaciones.

Si usas notebooks RSP, estás usando algo en lo que he estado trabajando. A fines de 2022 y principios de 2023, queríamos hacer que los laboratorios de usuarios en proceso bajo cargas pesadas fueran más confiables. Después de mucho esfuerzo, pude abordar todas las preocupaciones (fue un cambio arquitectónico importante) y lograr que el Kubespawner asíncrono fuera aceptado como la implementación estándar, por lo que ahora Jupyter lo mantiene, y la carga de soporte no recae únicamente en nosotros.

Restauro y juego con computadores y sistemas de videojuegos antiguos. La consola en funcionamiento más antigua que tengo es una Atari Super Pong de 1976; Tengo una colección bastante extensa de computadores y consolas de 8, 16, 32 y 64 bits desde la década de 1970 hasta la actualidad. Tengo tres perros grandes a los que adoro. Con entusiasmo, preparo y consumo cócteles artesanales, y recientemente he comenzado a andar en moto, pero no al mismo tiempo.

William O'Mullane

Data Management Project Manager, Deputy Project Manager (Software & IT)
Gerente del Proyecto de Gestión de Datos, Gerente Adjunto de Proyectos (Software y IT)
Joined April 2017



I keep Data Management (DM) and software in general running (fairly) smoothly. DM are responsible for getting the bits from the camera, turning them into science-ready products, and delivering them to the Science users. There is a big team for that, and I pretty much only do the management part--the excellent Technical Control Account Managers and product owners make it all work. This extends to Telescope and Site software and IT. I try to keep on top of all that is going on, support my managers in any way they need, smooth out issues, and resolve any contention between the various teams in the Rubin project. There are a lot of meetings to attend! Too many in fact, so I don't make it to all of them.

I am happy to have gotten and kept Data Management off the NSF problem list. But the biggest work achievement has been getting the Interim Data Facility okayed to run on a commercial cloud (Google). Thanks to the great team this has been very successful and can help transform the way the agencies see cloud computing.

I like to take photographs which are shared on flickr.com/womullan. Apart from reading the only hobby I have stuck with over the years is Kung Fu--though not competitively anymore.

En general, mantengo la Gestión de Datos (DM por sus siglas en inglés) y el software funcionando (bastante) bien. DM es responsable de obtener los bits de la cámara, convertirlos en productos listos para la ciencia y entregarlos a los usuarios de la ciencia. Hay un gran equipo para eso, y prácticamente solo me encargo de la gestión--los excelentes Administradores de Cuentas de Control Técnico y los propietarios de productos hacen que todo funcione. Esto se extiende al software de Telescopio y Sitio y a IT. Intento estar al tanto de todo lo que está sucediendo, apoyando a mis gerentes en lo que necesiten, suavizando los problemas y resolviendo cualquier discrepancia entre los diversos equipos en el proyecto Rubin. ¡Hay muchas reuniones a las que asistir! Demasiadas, de hecho, así que no llego a todas.
Me alegro mucho de haber sacado y mantenido a Gestión de Datos fuera de la lista de problemas de NSF. Pero el logro más interesante ha sido obtener la aprobación para ejecutar la Instalación de Datos Provisoria en una nube comercial (Google). Esto fue un gran éxito gracias al equipo fenomenal que ayudó a cambiar el punto de vista de las agencias sobre la informática en la nube.

Me gusta tomar fotos y subirlas en flickr.com/womullan. Leo muchos libros también, y con el paso de los años, sigo practicando artes marciales como el Kung Fu--pero ya no de forma competitiva.

Andrés A. Plazas Malagón

Associate Research Scholar | Académico Investigador Asociado (Universidad de Princeton)
Joined November 2018

I work in developing and implementing algorithms to calibrate the images that Rubin will take. I've also contributed to commissioning efforts, developing tools to take calibrations at the summit and automatically assess their quality. I'm also part of the LSST Dark Energy and Strong Lensing Science Collaborations. In DESC, I'm mostly active in the Sensor Anomalies and Weak Lensing Working Groups, and I've played leadership roles such as Chair of the Meetings Committee, member of the Membership Committee, and Liaison to Rubin Education and Public Outreach.

LSST will truly revolutionize astronomy, and I'm proud of the work that we do in the LSST calibrations team. I've been awarded grants by the LSST Corporation to hire an external organization to conduct an anti-Black racism workshop during our annual meeting, and to fund one of the first-ever paid REU-type programs for students in Colombian institutions. In recognition of my work, in 2016 I was awarded the National Prize in Exact, Physical, and Natural Sciences in my native Colombia, considered as the highest scientific honor in the country.

I love spending time with our 1-year old son, Sævar Arturo. I also enjoy flying single-engine planes, playing soccer, listening to heavy metal, going to concerts, and going on hikes and taking walks in our neighborhood park with my family, including our dog. Other hobbies I have include reading, playing video games, reading manga, watching anime, and learning languages (such as French, Japanese, and Russian).

Trabajo en el desarrollo e implementación de algoritmos para calibrar las imágenes que tomará Rubin. También he contribuido a los esfuerzos de puesta en servicio, desarrollando herramientas para tomar calibraciones en el cerro y evaluar automáticamente su calidad. Ademas, soy parte de LSST Dark Energy y Strong Lensing Science Collaborations. En DESC, participo principalmente en los Grupos de Trabajo de Anomalías de Sensores y Lentes Débiles, y me he desempeñado en roles de liderazgo como presidente del Comité de Reuniones, miembro del Comité de Membresía y soy el enlace de Rubin Education y Public Outreach.

LSST realmente revolucionará la astronomía, y estoy orgulloso del trabajo que hacemos en el equipo de calibraciones de LSST. He recibido subvenciones de la Corporación LSST para contratar a una organización externa para realizar un taller contra el racismo afroamericano durante nuestra reunión anual y para financiar uno de los primeros programas pagados tipo REU para estudiantes en instituciones colombianas. En reconocimiento a mi trabajo, en 2016 fui galardonado con el Premio Nacional de Ciencias Exactas, Físicas y Naturales en Colombia, considerado como el máximo galardón científico del país.

Me encanta pasar tiempo con nuestro hijo de 1 año, Sævar Arturo. También disfruto pilotear aviones de un solo motor, jugar fútbol, escuchar heavy metal, ir a conciertos, hacer senderismo y caminar en el parque de nuestro barrio con mi familia, incluido nuestro perro. Otros pasatiempos que tengo incluyen leer, jugar videojuegos, leer mangas, ver anime y aprender idiomas (como francés, japonés y ruso).

No more staff highlights apparently …

Status and Achievements

Data Facilities

- Most now have accounts on slack SLAC -
  - NCSA machines switched off
  - DAQ is in Chile - Base Test Stand (BTS)
- FRDF will do 40% processing
- UK still doing 25%
- Google contract signed for first 3 years of operations for US Data Access Center

Verification - we still have work to do

- LDM-692 gives a view of verification status - not much change in this year but a big set of tests are expected

	Priority	Fully Verified 2.1.1	Partially Verified 2.1.2	With Failures 2.1.3	Not Verified 2.1.4	Not Covered 2.1.5	Total
DM Requirements	(All)	124	87	1	328	237	**777**
LSE-61	1a	17	14		25	2	**58**
	1b	16	18	1	73	3	**111**
	2	3	4		44	3	**54**
	3				6		**6**
	Not Set					1	**1**
	(All)	36	36	1	148	9	**230**

DM summit ever closer to ops

- Multiple improvements to the Telemetry service (Sasquatch SQR-067)
  - Microsec latency to Engineering Facility Database (EFD)
  - EFD at SLAC with replication
- Prompt processing prototype set up on Google to allow rapid ad hoc processing of summit images. (DMTN-219)
- One square-degree as seen by the AuxTel
  - co-added image combines more than 2500 exposures in three different filters (SDSS gri) collected between February and May of 2022.
  - First time Rubin pipelines see data from a Rubin CCD.

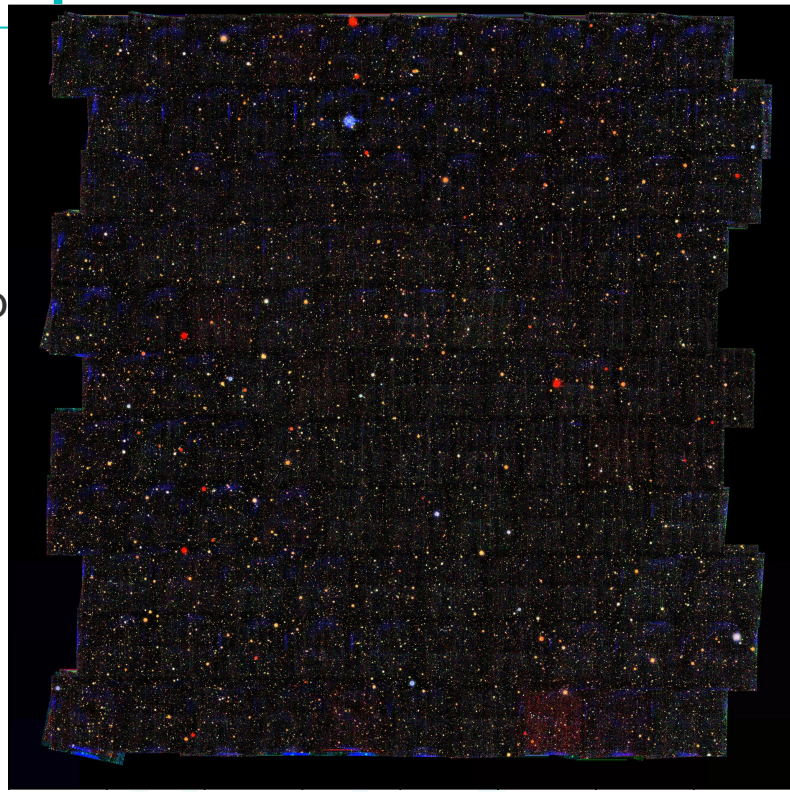


Figure credit: Erik Dennihy

Cisco ACI is gone !

Team photo after the Base ACI was powered off (yes, we used the big Enter button).

2023

- DP0.3 Solar system
- Summit databases

We also expect to spend a lot of time on commissioning and especially “missing functionality”

- FAFF - First Look Analysis Framework

Have a productive JTM

Thank you for the continued work on DM

Tomorrow we will get some relatively dark skies

Enjoy the summit - sunblock/lip balm

See how observing is working