Data Products

Melissa Graham, LSST DM Science Analyst, University of Washington

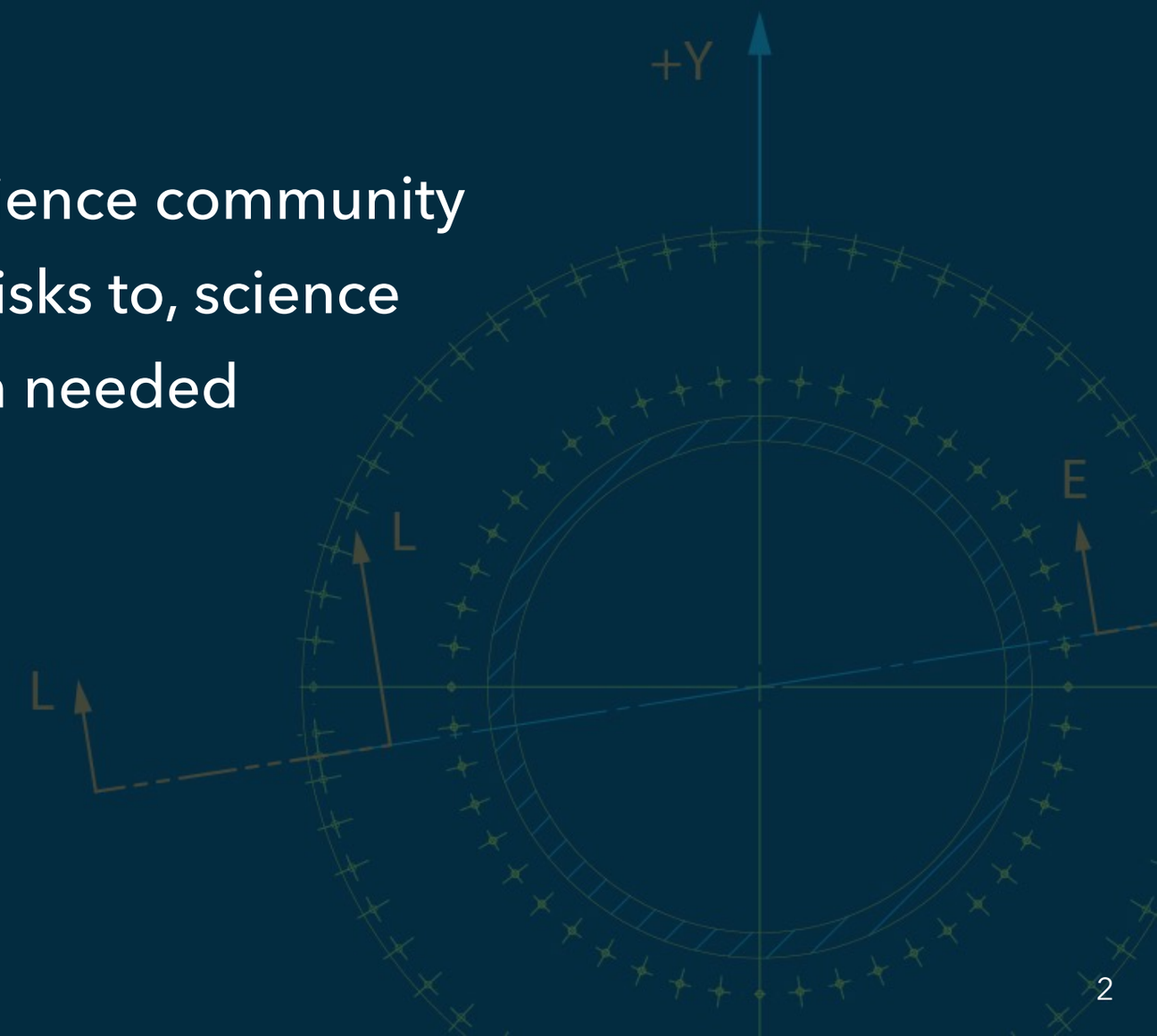
Large Synoptic Survey Telescope

Data Management System Science Team

Mandate: to ensure that the processing pipelines and data products will meet the LSST's science goals.

Method:

- understand the needs of the science community
- identify opportunities for, and risks to, science
- initiate change within DM when needed



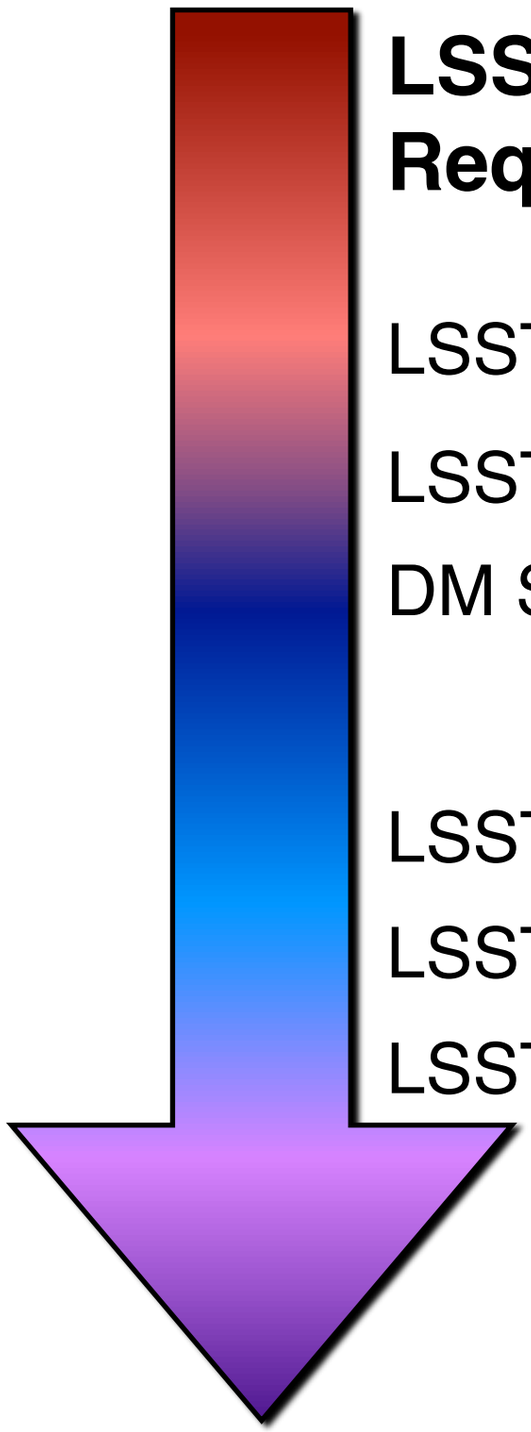
Supernovae!

The Progenitors of Type Ia Supernovae (SN Ia):

- occurrence rates and host galaxy stellar populations
  - *measuring the delay-time distribution*
- explosion mechanics from nebular phase spectra
- progenitor metallicity/density from early UV spectra
- evidence for non-degenerate companions
  - *time-varying narrow absorption from ionized CSM*
  - *hydrogen emission from swept-up material*
  - *late-time emission from an impacted companion*
  - *emission from SN ejecta interacting with CSM*

CSM = circumstellar material

LSST Data Products



LSST Data Management
Requirements Documentation "Flow-Down"

LSST Science Requirements Document (SRD) ls.st/lpm-17

LSST DM Subsystems Requirements (DMSR) ls.st/lse-61

DM Science Pipelines Design (DMSP) ls.st/ldm-151

LSST Data Products Definitions Document (DPDD) ls.st/lse-163

LSST Data Products Categories (DPC) ls.st/lse-291

LSST Science Platform Vision Document (SPVD) ls.st/lse-319

Everything in these slides regarding DM data products is from these documents.

LSST Data Products

Data Release Data Products
via Annual Data Releases

11 data releases in 10 years
Final database catalog: 15 PB

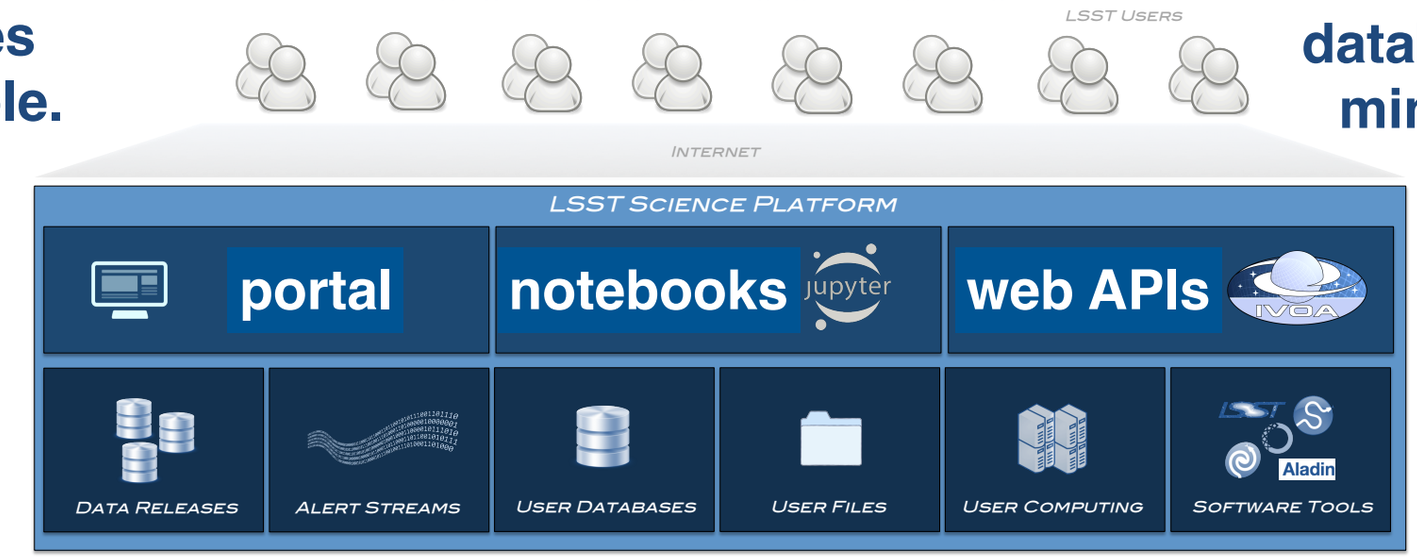
Prompt Data Products
via Alert Streams

Average ~10^6 per night
Real-time latency: 60 sec

Last 2 data releases available.

LSST Science Platform

Alerts database and mini-broker.



Data access via Data Access Centers & Services

LSST Data Products

Prompt

Previously "Level 1" data products

Real-time difference image analysis (DIA).

A stream of ~10^6 time-domain events per night (Alerts), detected, characterized, and distributed within 60 seconds.

A catalog of orbits for ~6 million bodies in the Solar System.

Data Release

Previously "Level 2" data products

Processed single-epoch and deep co-added images, and reprocessed DIA products.

A database of ~7×10^{12} detections (~30×10^{12} measurements) for ~37×10^9 objects (20×10^9 galaxies and 17×10^9 stars), produced annually and accessible online.

User Generated

Previously "Level 3" data products

User-produced added-value data products, e.g., deep KBO/NEO catalogs, variable star classifications, shear maps, etc.

Enabled by services and computing resources at the Data Access Centers and via the LSST Science Platform.

LSST Data Products

World Public

World Public data can be shared with anyone, with or without data rights.

Alerts: The full stream will be delivered to a limited set of community brokers who can share the Alerts with anyone.

Data Releases after 2 years: Could be accessed through collaboration with data rights holders, or by paying the “cost of shipping and handling”.

Education and Public Outreach: Limited data subsets for citizen science.

Proprietary

Proprietary data cannot be shared, and requires data rights.

Alerts Database: An archive of all issued Alerts.

Prompt Images and Catalogs: Difference images and source catalogs that are created and made available in real time (60s to 24h latency).

Annual Data Releases: Image stacks and source catalogs.

LSST Science Platform: Data portal, analysis toolkit, help-desk service, computational resources for user processing, an Alerts filtering service.

LSST Data Products

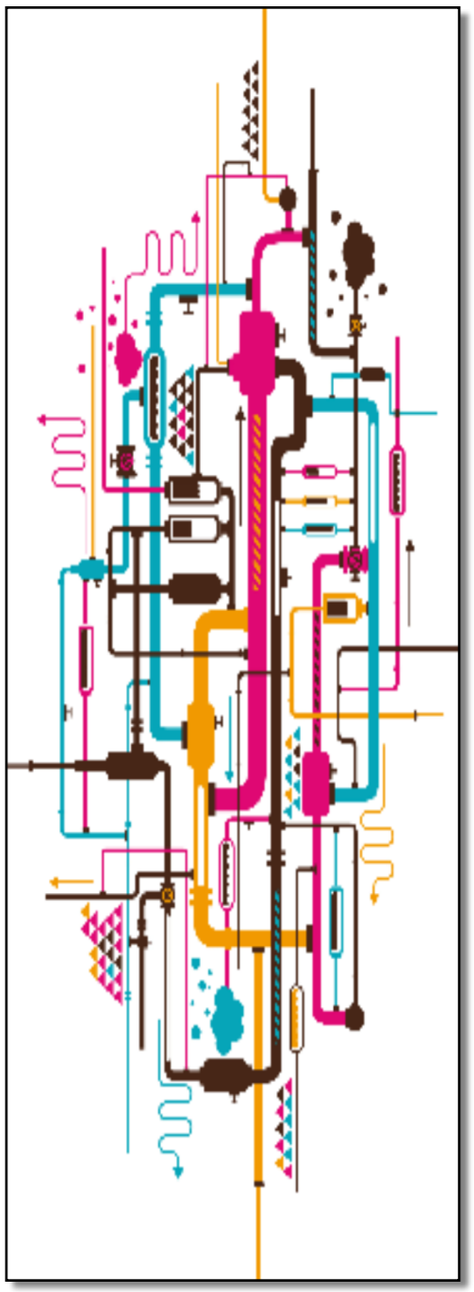
Definitions for Terms Commonly Used by LSST DM.

Standard Visit — an LSST observation (two 15 s ``snap" images)
Single Visit Image — reduced and combined standard visit
Template Image — a transient-free co-add of 6-12 months depth
Difference Image — subtraction of Single Visit from Template
CoAdd Image — a stack of images (i.e., median-combined)

DIA — Difference Image Analysis
DIASource — single detection on a single Difference Image
DIAObjects — association of DIASources, by coordinate
Alert — packet of information about a |SNR|>5 DIASource

DRP — Data Release Pipeline
Source — single detection in any image (Single Visit or CoAdd)
Object — association of Sources, by coordinate
ForcedSource — aperture photometry in a Single Visit Image at the location of an Object or DIAObject, regardless of its SNR

LSST Data Products



The Prompt Pipeline: Difference Imaging Analysis and Alert Generation

What happens within ~60 seconds of shutter close?

- Single Visit Image is processed and subtracted from the Template
- Difference Image detections with |S/N| > 5 become a DIASource
- DIASource characterization (PSF, flux, shape, etc.)
- DIASource association with existing DIAObject or SSObject,
  or the creation of a new DIAObject
- DIAObjects detected in past ~12 months get forced photometry
- DIAObject characterization parameters updated (e.g., variability)
- Alerts are issued (one per DIASource)
- DIA catalogs are updated in the US Data Access Center

What happens within ~24 hours?

- new DIAObjects get 30 day “pre-covery" forced photometry
- processed images become available in the US Data Access Center
- Moving Object Pipeline Software (MOPS) runs on DIASource catalog

LSST Data Products

The Prompt Pipeline:
Difference Imaging Analysis and Alert Generation

What kinds of measurements are in the DIA catalogs?

DIASource
- coordinates, and association with DIAObject or SSObject
- time of mid-exposure at its location on the CCD
- flux in the difference and visit images (PSF, aperture)
- shape parameters (trails, dipoles, FWHM, extendedness)
- parent/child deblending flags

DIAObject (association of DIASources)
- time-averaged coordinates; parallax & proper motion
- fluxes by filter, time-averaged and single-visit
- periodic and non-periodic variability features
- association with Data Release object catalog

LSST Data Products

The Prompt Pipeline:
Difference Imaging Analysis and Alert Generation

What is an `Alert` packet?

- text file containing schema and data for 1 DIASource
  ↳ *VOEvent format or similar*
- full record of the triggering DIASource
- entire associated DIAObject or SSObject records
- **last 12 months** of DIASource records
- matching Object IDs from latest Data Release Object image stamps
- image stamps
  ↳ *at least 6”x6”; difference and template; flux, variance, and mask; includes meta-data such as WCS, zero-point, PSF*

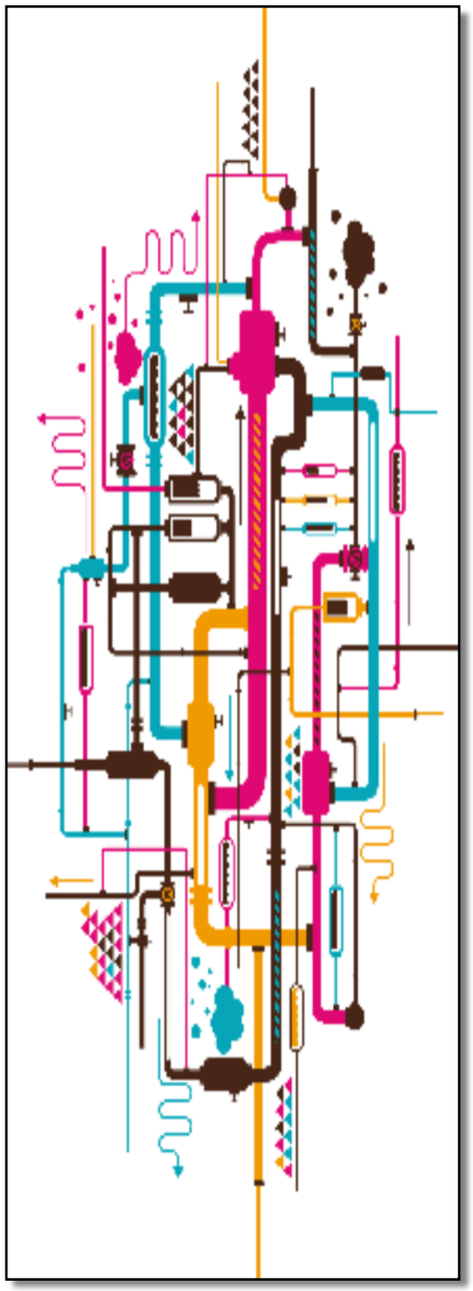
What is an Alert Broker?

Alerts will be delivered to a limit set of community brokers that filter and classify events to enable scientific use.

↳ LSST will provide a limited capacity filtering service, a “mini-broker”, in the Science Platform.

LSST Data Products

The Data Release Pipeline



What kinds of images are produced?

- raw and processed Single Visit images
- stacked images (CoAdds):
  - short-period (e.g., yearly)* and full survey
  - best seeing and deepest (unless equivalent)
  - for each filter *ugrizy*, and a multi-color*
- transient-free template images

**not persisted, but used for measurements in the catalogs; can be recreated by users*

What kinds of measurements are in the DRP catalogs?

Source (and **ForcedSource**)

- coordinates, fluxes in single-visit images
- deblending (parent/child identifiers)
- model fits (bulge/disk, exponential, petrosian, kron)
- surface brightness, extendedness parameters
- color (seeing-independent)

Object (association of Sources)

- similar characterization parameters as Source
- photo-z; variability characterization
- proper motions and parallaxes

Database Schema Browser: https://lsst-web.ncsa.illinois.edu/schema/index.php?sVer=baseline

LSST Data Products

The Data Release Pipeline

Each annual data release will also contain a full reprocessing of all images with the DIA pipeline.

There are 2 versions of DIA products.

To meet different science needs:

Prompt ("living" data products)
- updated in real time with every readout
- contains last ~12 months of data
- variability parameters based on past 12 months
- associations with the "Yearly" DIA catalog products

e.g., real-time follow-up

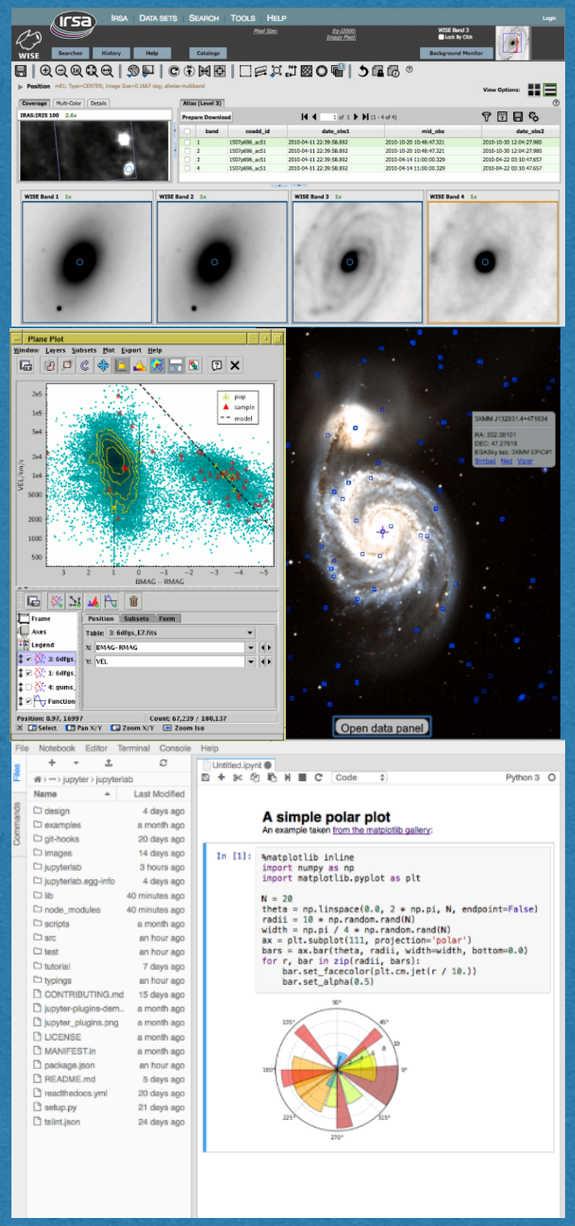
potted flowers
vs
cut flowers

Data Release
- full LSST data set reprocessed with latest codes
- variability parameters based on full survey to date
- forced photometry for all objects in all epochs
- same types of products as the "Living" versions

e.g., rates, population studies

LSST Data Products

User-facing tools will look familiar.

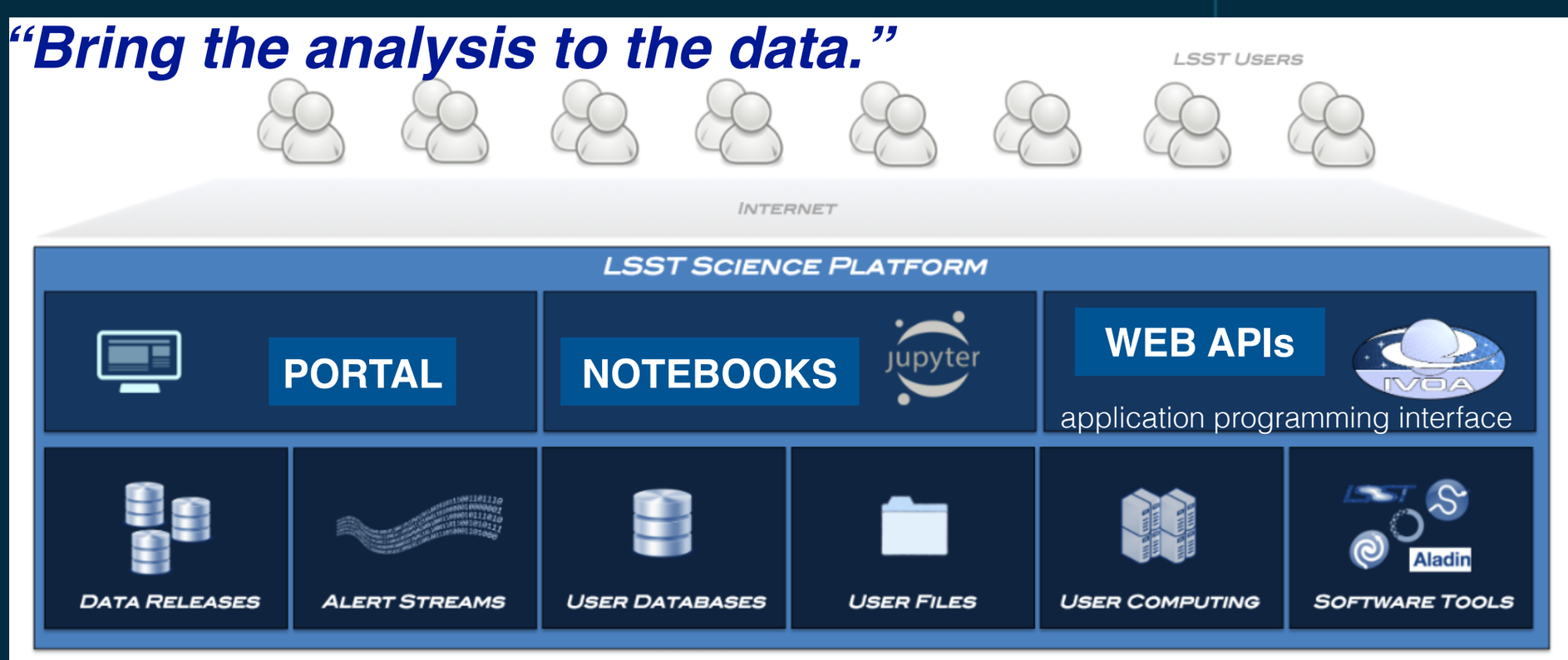


What is the LSST Science Platform?

A toolkit, a workspace, and a portal to the data.

- no need to download the data products
- tools for browsing, visualization, analysis
- use of pre-installed code libraries
- computational resources for query/processing
- view demos at http://ls.st/bgt

"Bring the analysis to the data."



LSST Special Programs

With regards to Special Programs (SP), LSST Data Management:

- ***will not*** write unique algorithms for processing SP data
- will incorporate SP data into the Prompt and DR products when scientifically beneficial (e.g., Alerts, CoAdds, catalogs)
  - *with program of origin included as a flag for queries*
  - *Alerts can be issued from SP that use Standard Visits*
- will reconfigure its pipelines to generate separate imaging and catalog products for SP data, whenever possible
  - *SP data may be released on intermediate timescales*
- will make the Software Stack source code available to the community
- will allocate an *additional* ~10% of the LSST computing resources for user-driven analysis and data product creation in the US DAC

LSST Special Programs

Software Constraints in the Call for White Papers (Section B.1)

"The Project will not take formal responsibility for specialized data reduction algorithms needed to process data."

Specialized data reduction algorithms might include:

- processing for non-standard visits
  - *e.g., short exposures in which the PSF is not well-formed*
- source deblending in extremely crowded fields
- shift-and-stack for faint moving objects
- computational needs surpassing that available in the Science Platform

"If a proposed dataset will require special processing, a plan to obtain the necessary software and compute resources must be provided in the white paper."

LSST Special Programs: A Timeline

First four extragalactic deep fields chosen.

First round of white paper proposals for LSST Special Programs.

planning continues through Science Collaborations, Observing Strategy White Paper

October 2018
Deadline for white paper proposals

June 2018
Call for white paper proposals

DMTN-065

Project Science Team defines criteria; Science Advisory Council reviews proposals and makes recommendations.

Latest set of Special Programs incorporated into the Operations Simulator.

planning continues

2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022

DMTN-065: “Data Management and Special Programs”, assesses DM’s plans for processing the diversity of raw data that may be generated by the community’s Special Programs proposals.

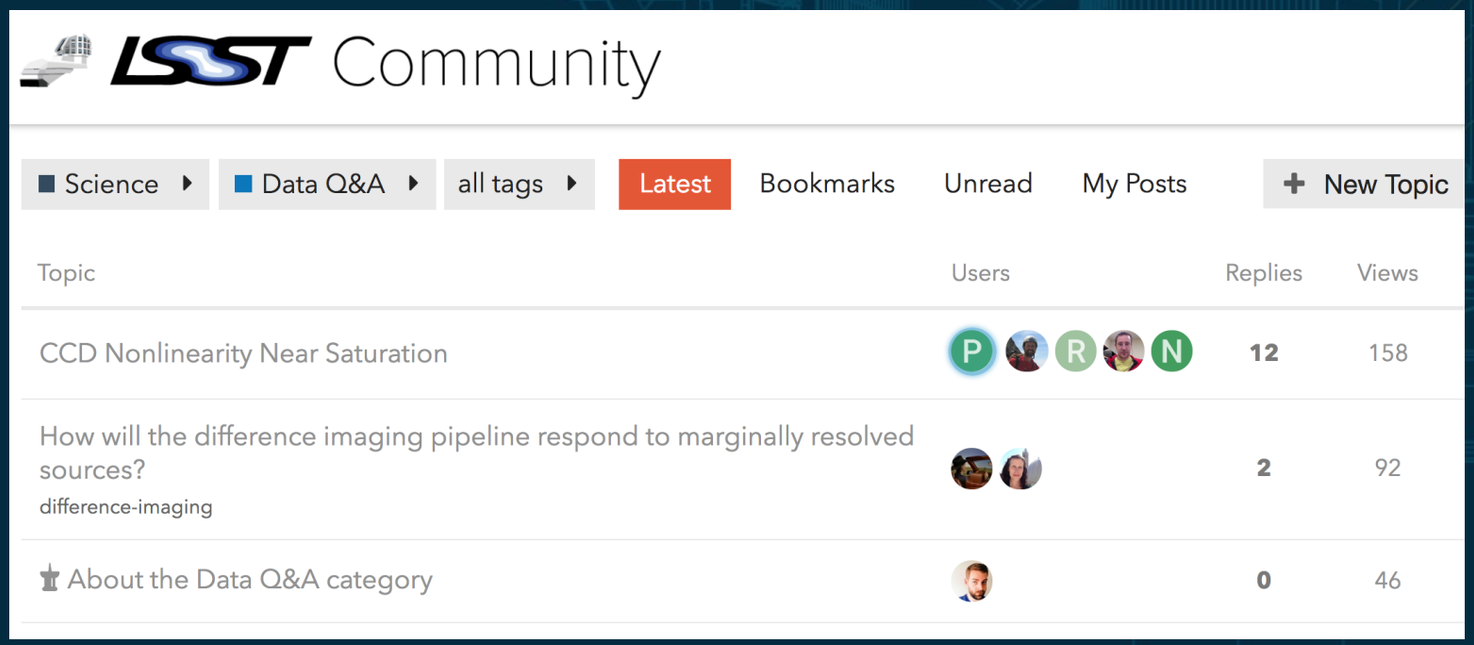
Thank you for listening.
Thanks to Federica and Rachel for organizing.

Questions?

Feel free to contact me: mlg3k@uw.edu

Or, use the Community forum “Data Q&A”.

The DM team monitors all threads and provides timely answers.



Additional Links for LSST Special Programs

LSST Project Website project.lsst.org
Observing Strategy White Paper github.com/LSSTScienceCollaborations/ObservingStrategy

LSST DDF Information https://www.lsst.org/scientists/survey-design/ddf
Deep Drilling Fields White Papers https://project.lsst.org/content/whitepapers32012
The 2016 LSST Project and Community Workshop project.lsst.org/meetings/lsst2016 (slides available)
"General Review of Proposed DDF/MS", Neil Brandt, LSST AHM, Aug 2016:
https://project.lsst.org/meetings/lsst2016/sites/lsst.org.meetings.lsst2016/files/Brandt-DDF-MiniSurveys-01.pdf
"Simulations, Metrics, and Merit Functions for Mini-Surveys and DDF", Steve Ridgway, LSST AHM, Aug 2016:
https://project.lsst.org/meetings/lsst2016/sites/lsst.org.meetings.lsst2016/files/Ridgway-SimulationsMetrics_1.pdf
"Draft thoughts on selecting DDFs", Beth Willman, LSST SAC Meeting, October 2016:
https://project.lsst.org/groups/sac/sites/lsst.org.groups.sac/files/Willman_DDF.pdf

Community forum on DM concerns for Special Programs:
https://community.lsst.org/t/deep-drilling-fields-and-data-management/1115

Data Management and LSST Special Programs Document: http://ls.st/dmtn-065