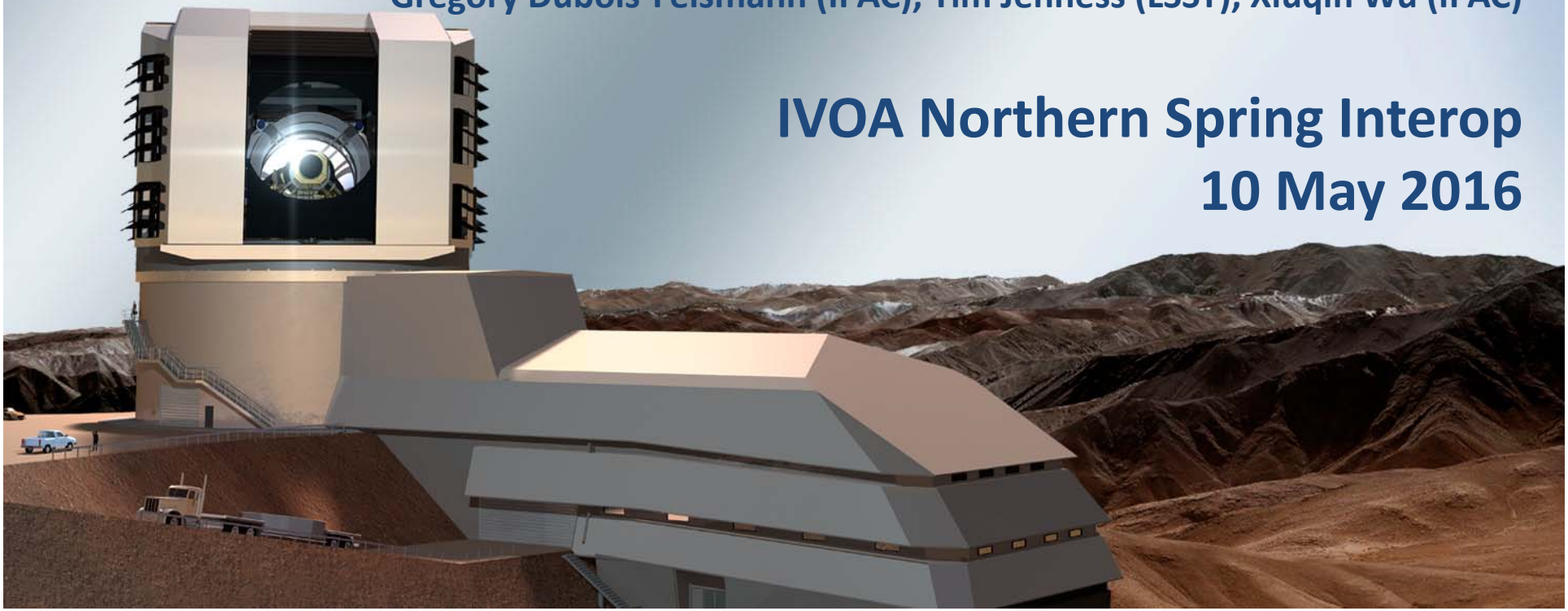
Large Synoptic Survey Telescope and Synergies with the VO

David R. Ciardi (IPAC/Caltech)

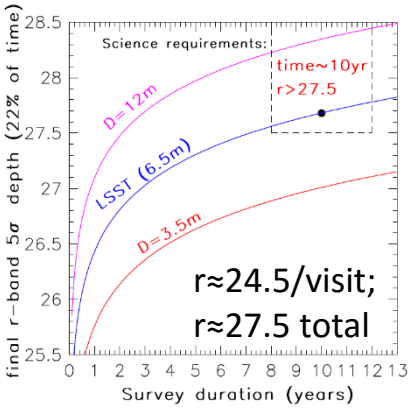
John Swinbank (Princeton), Brian van Klaveren (SLAC)
Gregory Dubois-Felsmann (IPAC), Tim Jenness (LSST), Xiuqin Wu (IPAC)

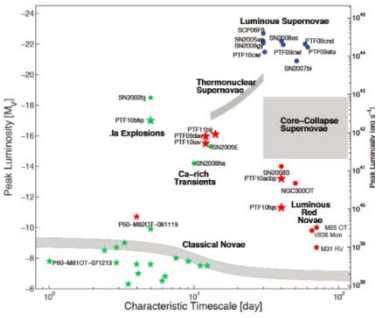
IVOA Northern Spring Interop
10 May 2016

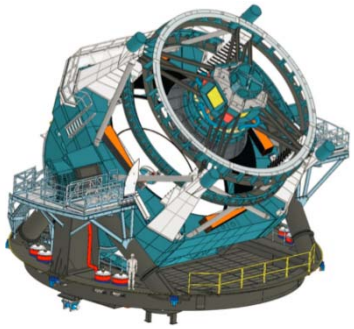


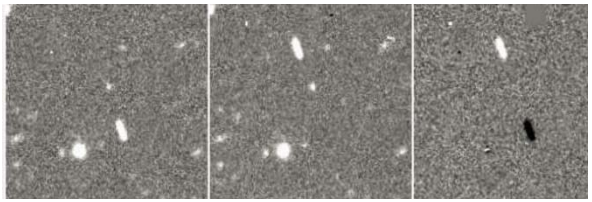
LSST: In One Slide

- 10 year Deep, Wide Fast survey
  - 18000+ deg2, 800 – 900 visits per position on sky
  - 2x15s back to back exposures
  - ~10% of available for other projects (e.g., deep drilling)
- 8.4m telescope and 9.6 deg^2 Camera
  - 189 CCDs; 3.2 giga-pixel
  - 0.2'' per pixel
  - 6 bands (ugrizy; 320-1050nm)
- Time domain science:
  - Novae, black-hole binaries, GRBs …
  - Source characterization
  - Instantaneous discovery
- Census of the Solar System:
  - NEOs, PHAs, moving objects
  - Solar system & planet formation
- Mapping the Milky Way:
  - Structure and accretion history
  - Properties of all stars within 300 pc
- Dark energy and dark matter:
  - Strong Lensing
  - Weak Lensing
  - Supernovae

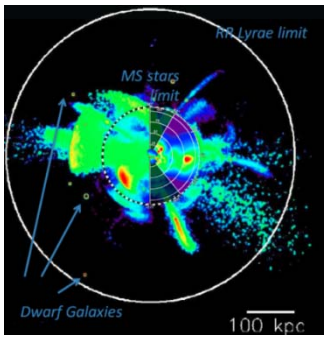


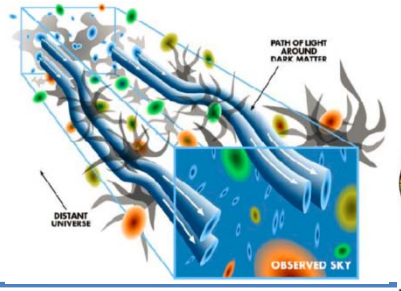


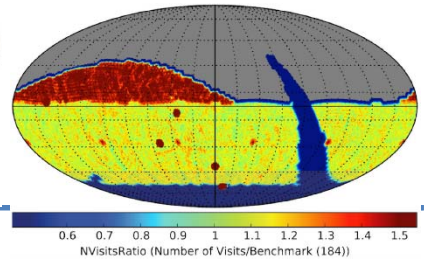




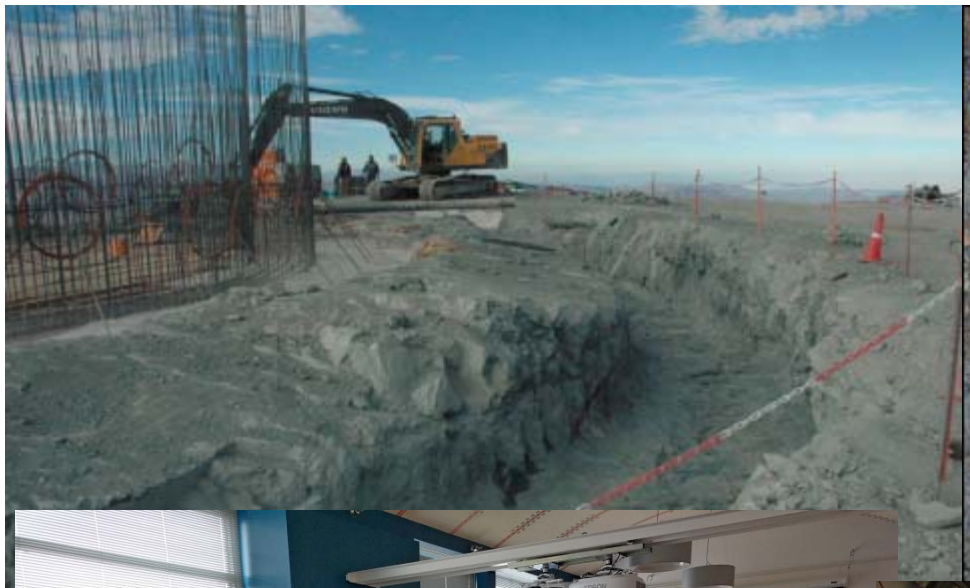
- Construction: 2014 – 2019
- Commissioning: 2019 – 2020
- Science Verification: 2021 – 2022
- Operational Survey: 2022 – 2032





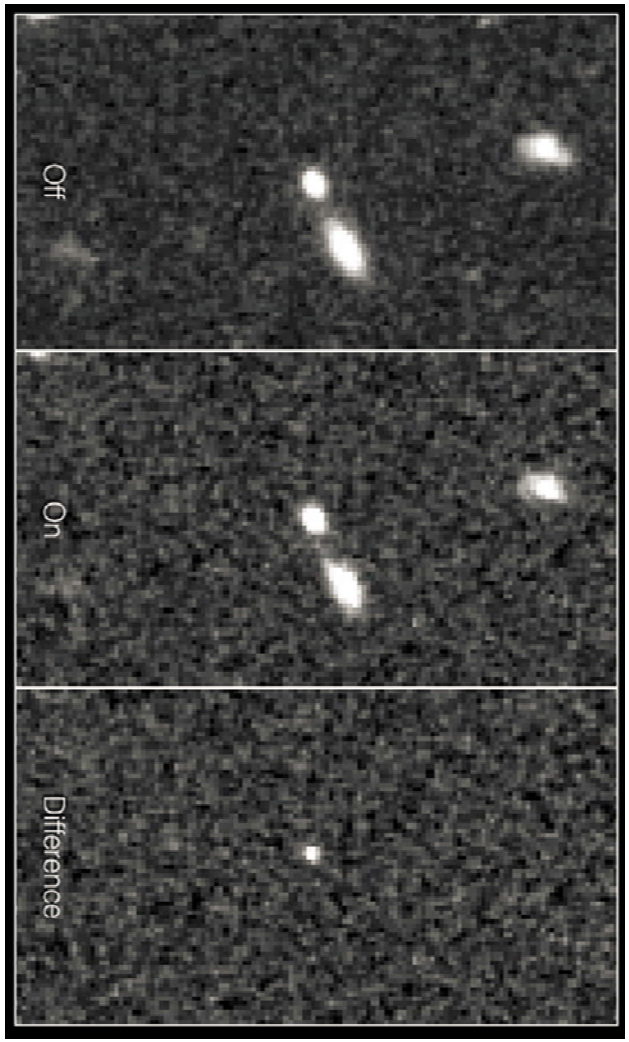


Real Construction



Level 1: Alerts

- Alerts issued within 60 seconds of observation
  - 10,000,000/night (average), 10,000/visit (average), 40,000/visit (peak)
- Each alert includes:
  - Position
  - Flux, size, and shape
  - Light curves in all bands
  - Variability characterization (e.g., low-order light-curve moments, probability the object is variable, previous id)
- Cut-outs centered on the object (template, difference image)
  - Including appropriate metadata



Level 2: Annual Releases

- Calibrated and consistent catalogues & images
  - Objects, detections, detections in difference images, coadds, etc
  - Enable static sky science and time-domain science which is not time sensitive (e.g. statistical investigations of variability)
  - Made available in annual Data Releases
  - Two releases in the first year
- Complete reprocessing for each release
  - Every DR will reprocess all data
  - Including reprocessing of level 1 data
  - Forced photometry
- Projected catalog sizes:
  - 18 billion (DR1) ➜ 37 billion (DR11) separate objects
  - 750 billion (DR1) ➜ 30 trillion (DR11) individual measurements
  - Cumulative ~500 PB image and ~50 PB catalogue data

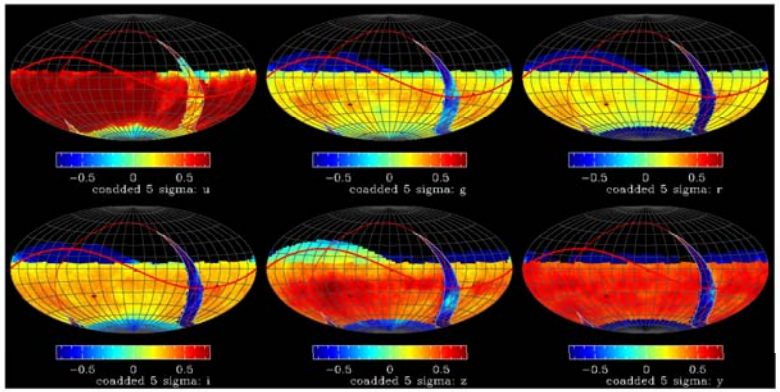
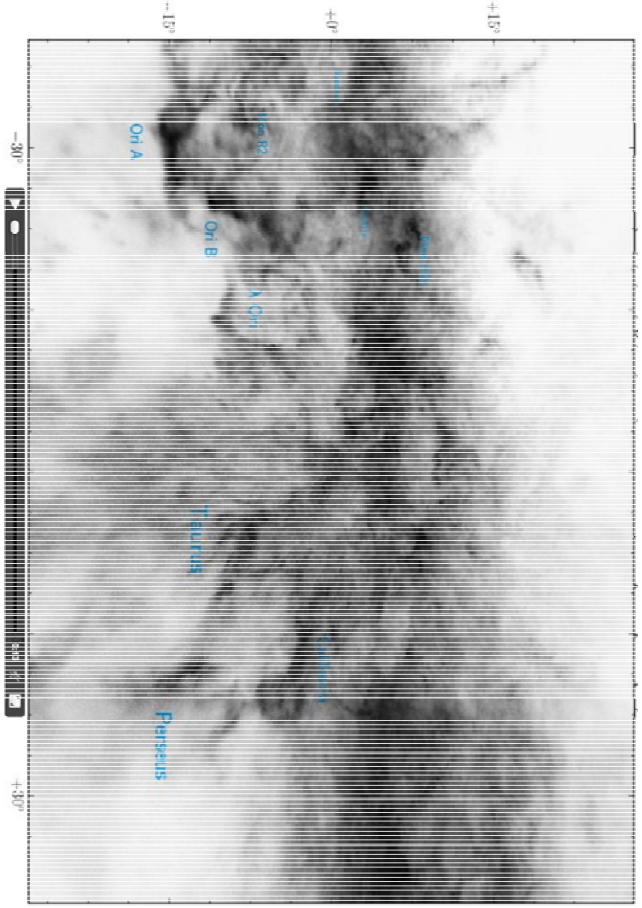


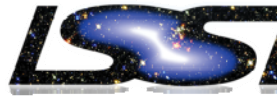
Figure 3.2: The 5σ stacked point-source depth of the simulated ten-year survey shown in Figure 3.1. The scale in each panel shows the depth of the stack relative to the fiducial values of 25.8, 27.0, 27.2, 27.0, 25.7, and 24.4 in u, g, r, i, z, y respectively.

Level 3 and Workspace Environment

- Products created by the community and made available both privately and publicly
- Use-cases not fully enabled by Level 1 and 2:
  - Reprocessing images to search for SNe light echos
  - Characterization of diffuse structures (e.g., ISM)
  - Extremely crowded field photometry (e.g., globular clusters)
  - Custom measurement algorithms
- Enabling Level 3:
  - User databases and workspaces
  - Non-LSST data: catalogs and files (e.g., images, spectra, time series)
  - Enabling user computing next to the data
  - Sized for ~10% of total compute budget
  - Making the LSST software stack available to end-users

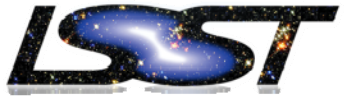


LSST Science User Interface and Tools



- The SUIT is the entry point of the community to the LSST data – for expert and novice users
  - The SUIT needs to be simple enough to engage the novice and flexible enough to meet the needs of the power-users
- The SUIT intended to be a toolkit
  - No way we can anticipate all the needs/wants of the community
  - Enable creativity and flexibility
- Build a portal that fulfills the needs of the general user (e.g., searching, image visualization, table manipulation, plotting, workspace etc.)
- Provide a python workspace environment that is connected to the web-based portal and the working environment underneath
- Components will be usable by others to use and to build tools that meet their own special needs

LSST Science User Interface and Tools



- Visualization tools server-based with thin client to enable large dataset interaction
- User workspace environment with access to computation resources at data access center(s)
- Remote client access to environment
  - Data access and visualization APIs (with python wrappers)
  - Web browser
  - Jupyter notebook
- Enable the ability to connect local resources to LSST DAC resources
- Enable the ability to install the SUIT and workspace environment locally
- Enable users to create their own environments for their own purposes with their own tools
  - If you don't want to use the visualization/interactive tools developed for the portal (and usable within the python environement), you will be able to build and use your tools/enviroment
  - Interactive visualization controllable by JavaScript and Python API
  - Customizable visualization functions through call-backs provided by users

VO Areas of Consideration

- LSST is working on how best to provide access to the data
  - How to best take advantage of the VO work
  - Learn from, as well as contribute to, other projects
- SCS, SIAP, and TAP access will (almost certainly) be enabled within the interfaces and workspace
  - Started on a Python (thin-layer) TAP service in front of database
- CAOM
  - Model could help make footprint services easier to implement
  - Currently does not handle non-image data (e.g., catalogs)
  - Considering making tables CAOM capable or having a CAOM-compatible view
- Datalink
  - A given query has potential to return hundreds to thousands of images and data
  - Linking images – sources – time series (one-to-many and many-to-one) is crucial
- VOSpace
  - Need a workspace environment for users to store results, run code, share results, Level-3 data
  - Workspace science and technical requirements still under development
  - As this becomes more defined, we will revisit the effectiveness of utilizing VOSpace

Time Series Data Model Under Consideration

- LSST Time Series is multi-dimensional and multi-type
  - Photometry (multi-filter)
  - Images (e.g., calibrated, difference)
  - Analysis (e.g., psf Model, shape fitting)
- Simple Time Series (STS)
  - Might be sufficient for simple photometry time series but may not be able to convey all the needed and appropriate metadata (see IVOA Note from June 2015 comparing STS and SSAP for OMC archive).
  - Not flexible enough to convey all of LSST data complexity
- Simple Spectral Access Protocol (SSAP)
  - May be better suited for photometry time series (see IVOA Note from June 2015 comparing STS and SSAP for OMC archive).
  - But is currently limited to 1-d spectra and may not be able to handle multi-dimensional and/or multiple types of data

VOEvent: Under Consideration

- Concerns
  - XML Packet is verbose with a large amount of overhead per event
  - Packet format is not ideal for binary data (e.g., images)
  - Ensuring the authenticity and scientfic integrity of the stream
  - Is transfer protocol sufficient for large number and rate of events?
  - Repeated information … many alerts per image and all from the same telescope

- We have a funded LSST Corporation proposal to engage the community about how to best evolve
  - VOEvent and VOEvent Transport Protocol
  - Annotations
  - More complex data
  - Needs to feed brokers that are being developed (e.g., Antares) as well as internally developed broker

Alert System Questions

- Science Needs of the community
  - How are unique events identified and labeled?
  - How are events characterized?
  - How are events presented in a simple and digestible form?
  - How are multi-wavelength, multi-dimensional time-series information packaged?
  - How are annotations and additional data, necessary for the characterization and classification of events, supported?
  - What is the scientific workflow of event follow-up and how does the alert system support such a workflow?
- Current Infrastructure
  - Can the current event infrastructure be adapted to address the scientific needs of the community?
  - Is the current alert distribution infrastructure scalable to the needs of LSST?
  - What infrastructure is missing or poorly defined?
- What do the brokering communities actually need?
  - Many current dedicated examples: supernovae community, microlensing community, NEO community,

Time Scale for Time Series/Alert System Study

- Need to engage other projects (e.g., PTF/ZTF, SKA, LIGO, LOFAR), but also engage the consumers of the time series and alerts (e.g., SN, NEO, microlensing)
- Would like to establish a set of telecons over the next three months
  - Concentrate on Alert issues as this may need more community input on needs
    - Telecon 1 (June): Layout of the problem and plan for quantitative discussion
    - Telecon 2 (early-July): Project oriented issues
    - Telecon 3 (late-July): Consumer oriented issues
  - Time series data model may be best 'explored' by building against real data
    - LSST has some internal build milestones against real data over the next 6 - 12 months that may prove useful in identifying areas
- Aim for 'draft' reports
  - "Northern Autumn" = "Southern Spring" Interop Meeting for alerts
  - "Northern Spring" = "Southern Autumn" Interop Meeting for time series data model