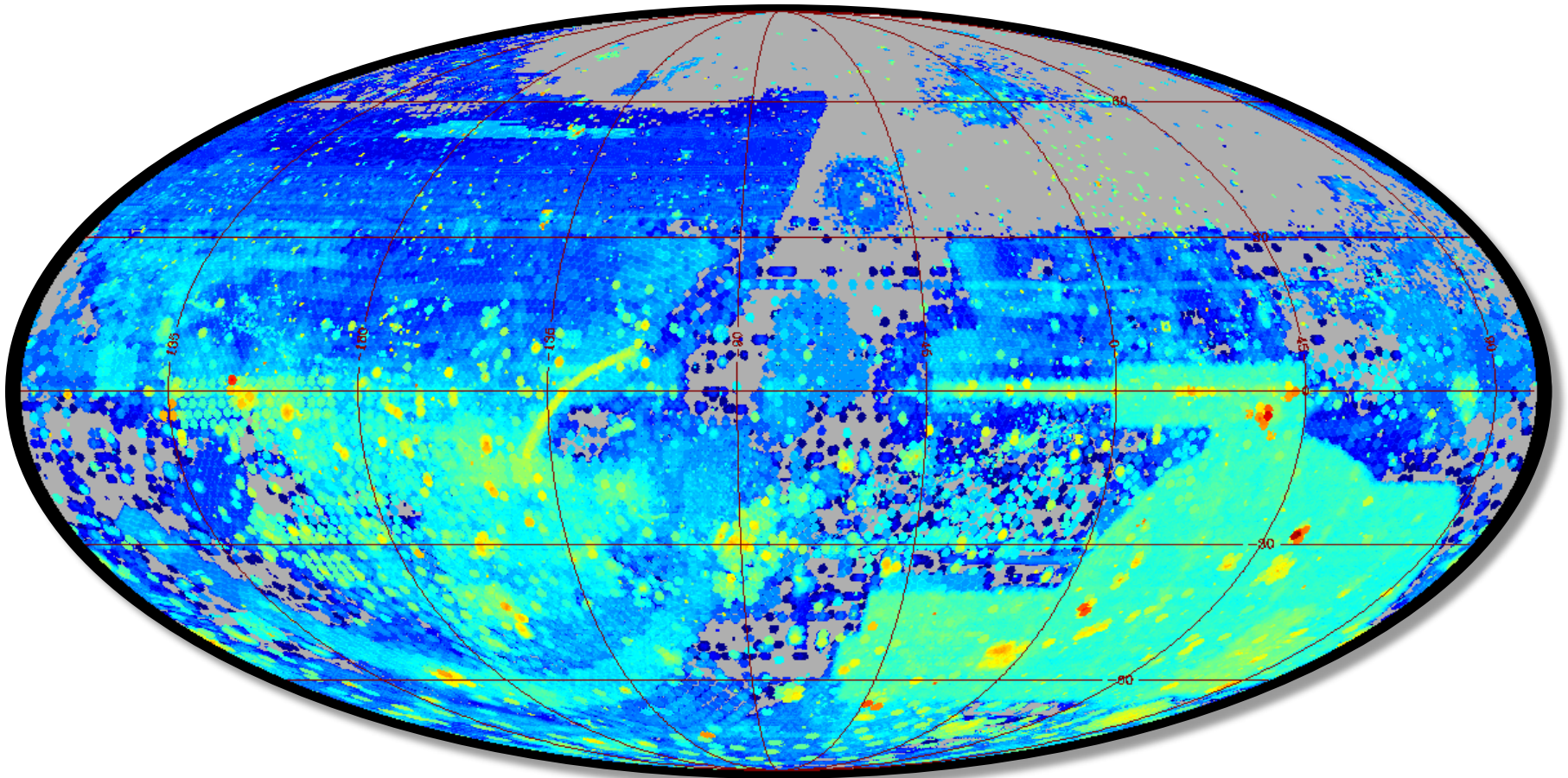


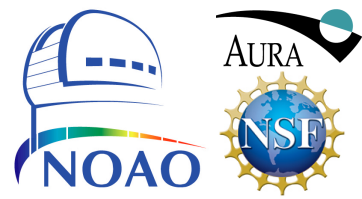


Community Science and Data Center

Adam Bolton, Associate Director

Dara Norman, Deputy Associate Director





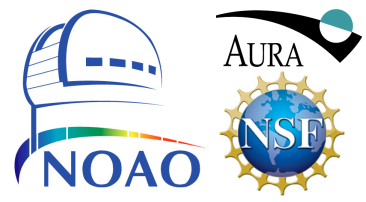
CSDC and the NOAO Mission

The “third mountaintop”

NOAO’s mission is to provide access to forefront astronomy research capabilities without regard to institutional or collaborative affiliation.

CSDC supports this mission through

- **Development and operation of capabilities for data-intensive astronomy**
- **Coordination across the US OIR System of ground-based observing facilities**
- **Community development initiatives**



Data Intensive Astronomy



CSDC and the Evolving NOAO Mission

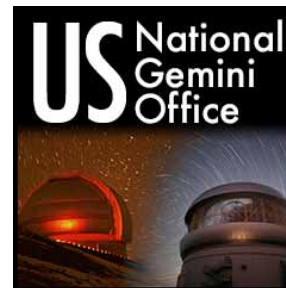
Open skies, open data

Traditional measure of opportunity



NOAO UC, Tucson, May 2017 (D3)

NOAO open-access science platforms

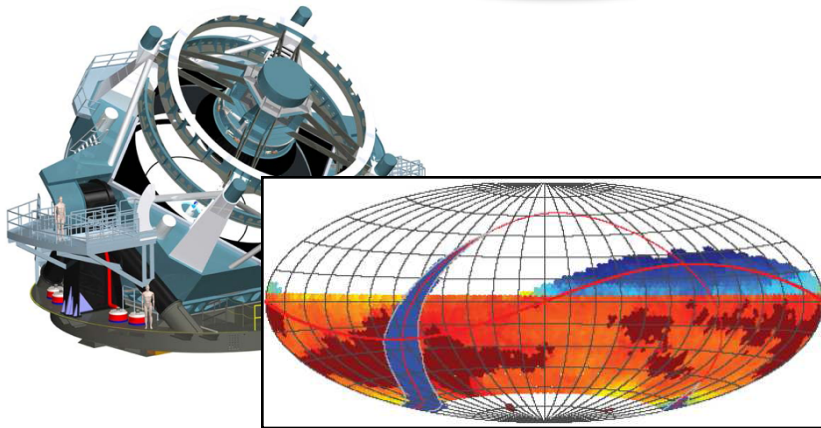
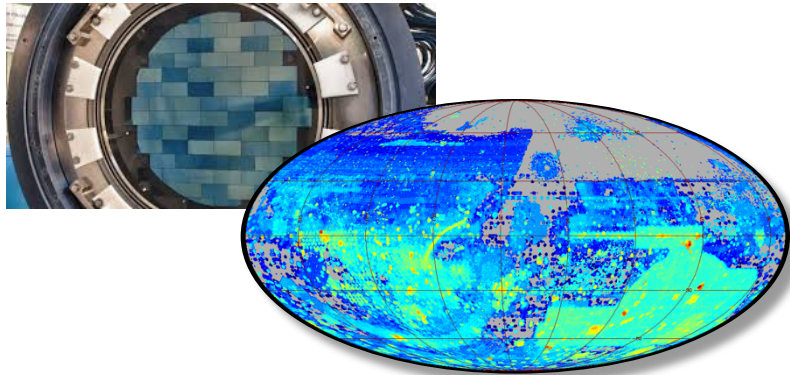




CSDC and the Evolving NOAO Mission

Open skies, open data

**Big-surveys/big-data
measure of opportunity**



**NOAO open-access
science platforms**





Major Survey Datasets at NOAO

Current imaging surveys

Survey	Camera	Details
DES*	Blanco / DECam	Cornerstone project for Dark Energy Camera: 5-year survey (2013-2018) covering 5,200 deg ² in <i>grizY</i> to $z=24.7$. DR1 covering years 1-3 (coadd images and catalogs) over full footprint scheduled for public release through NOAO and NCSA in December 2017.
DECaLS*	Blanco / DECam	$0 < \text{Dec} < +30^\circ$ DESI targeting survey: 6,700 deg ² in <i>grz</i> to $z=23.0$. Processed by Community Pipeline and “Tractor” code. DR3 with 4,200 deg ² of 3-band coverage available now at NOAO (images and catalogs).
SMASH	Blanco / DECam	480 deg ² Magellanic cloud coverage in <i>ugriz</i> to $z=23.5$. 61/197 fields (images & catalogs) now public at NOAO.
MzLS*	Mayall / Mosaic3	5,000 deg ² DESI targeting survey to $z=23.0$. 3,000 deg ² images & catalogs coming soon with DECaLS DR4.
BASS*	Bok / 90prime	5,000 deg ² DESI targeting survey in <i>gr</i> to $r=23.6$. 3,000 deg ² images & catalogs coming soon w/ DECaLS DR4.

***Contributes to 14,000 deg² total DESI targeting footprint**

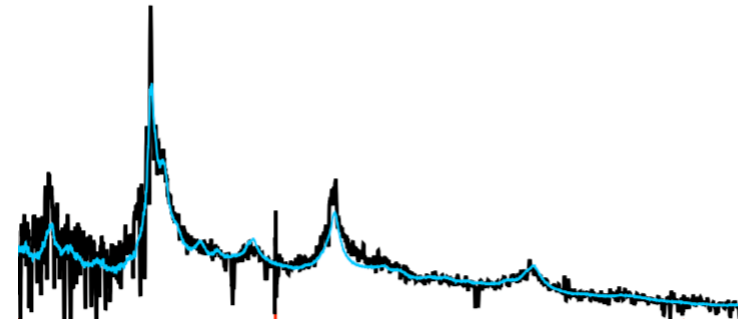


Major Survey Datasets at NOAO

Spectroscopic surveys: SDSS and DESI

- **Sloan Digital Sky Survey (SDSS)**

- Operating 2000-2020, currently in fourth phase (SDSS-IV)
- Over 4 million spectra collected to date
- NSF supplemental funding awarded to host mirror copy of SDSS data archive at NOAO for integration with Data Lab (in progress)
- Enable co-analysis with DES/DECaLS/MzLS/BASS
- Develop survey-scale spectroscopy hosting capability for DESI



- **Dark Energy Spectroscopic Instrument (DESI)**

- 5,000-fiber spectrograph being deployed to KPNO Mayall 4m
- 5-year spectroscopic survey of over 40 million objects
- NOAO plans to be the primary public release point
- NOAO working with DESI project and NSF to maximize astronomical community benefit of the DESI survey data set



Motivation: data-intensive archival science

The scientific opportunities of large survey data sets at NOAO require new capabilities for discovery, exploration, query, access, and analysis



Data Lab is NOAO's initiative to deploy these capabilities in order to **enable community users to realize these data-intensive science opportunities**

Data Lab also aims to provide users with a platform to develop transferrable skills that will **enable future research with the LSST data set**



NOAO Data Lab

Current feature overview

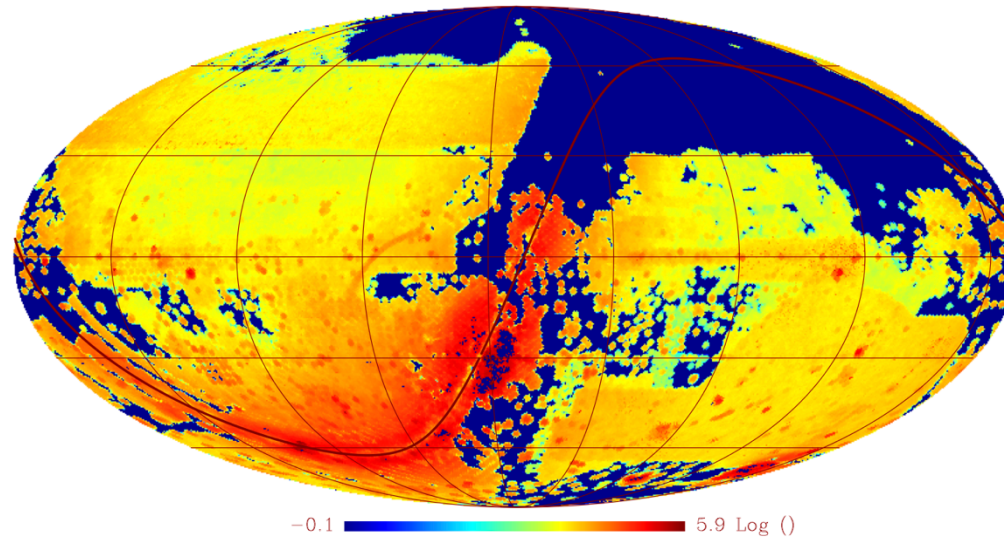
With public release in June 2017, Data Lab will provide users with

- Services for querying large catalogs at NOAO
- Access to cutouts from NOAO archival image data
- Virtual user storage co-located with the NOAO archive
- Jupyter IPython notebook support, with examples
- Command-line and API access to Data Lab services
- Online documentation for all of the above

*DECaLS, SMASH, and other data sets available now;
DES DR1 coming in December 2017!*

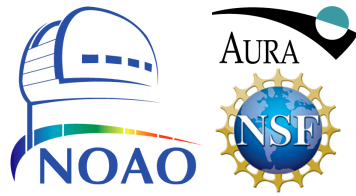
```
In [2]: # query DB for a list of all SMASH fields
query = "SELECT ra,dec,nexp,nobj FROM smash_dr1.field WHERE nexp IS NOT NUL
result = queryClient.query(token,adql=query,fmt='csv') # submit the query,
print "Inspect the returned result: it's a csv-formatted string:"
print result[:105]
```

Publicity through multiple tutorials, demonstrations, and hack sessions



- NOAO “all sky catalog” of detections and objects across entire NOAO archive
- “Containerization” of user-defined workflows for execution on Data Lab systems
- Addition of other survey catalogs and image data sets
- Data publication capabilities

See Olsen/Juneau/Nikutta presentation tomorrow for more details

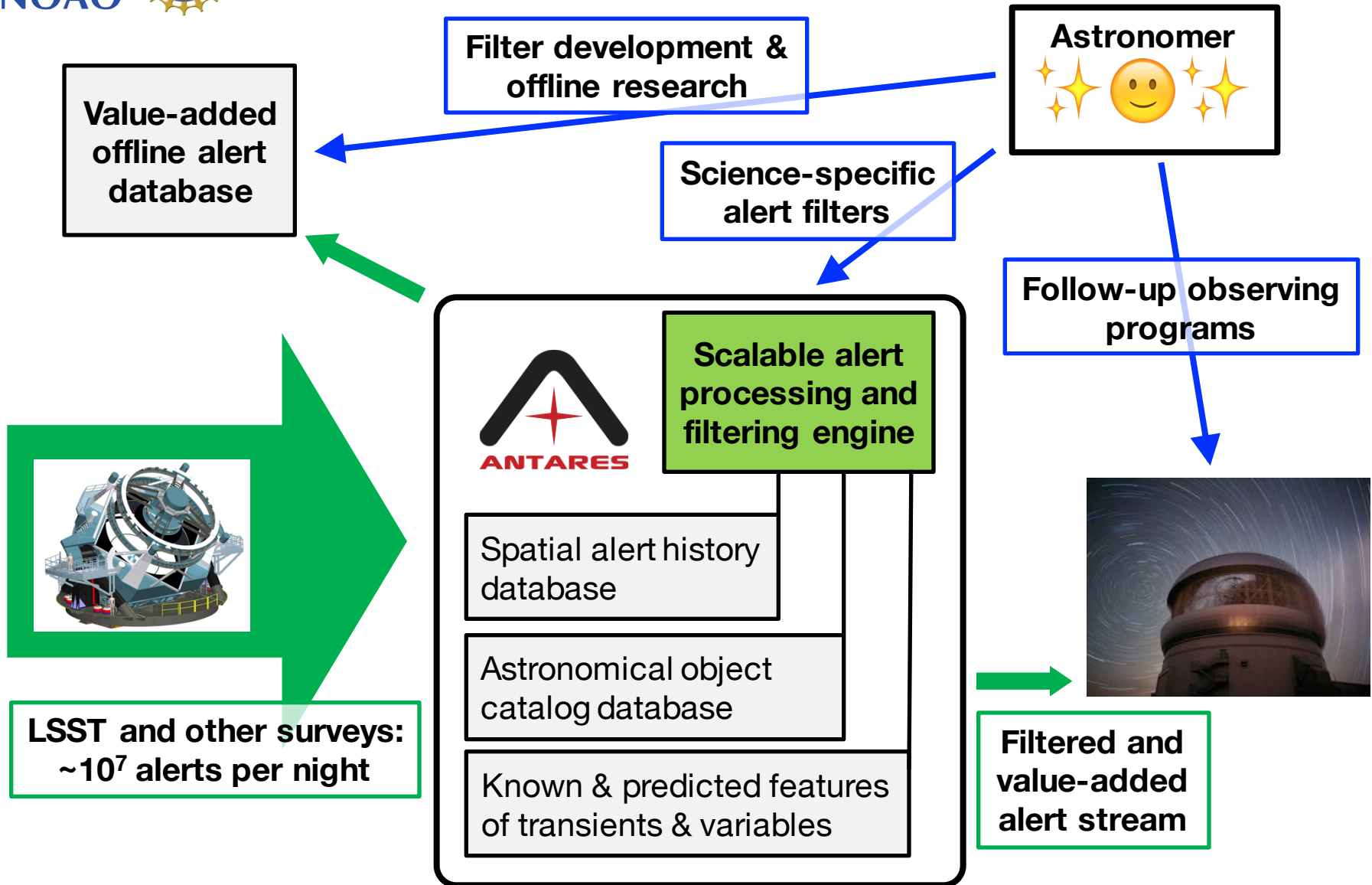


- ANTARES goal: develop, test, and deploy a flexible public platform for implementing time-domain science programs at the scale of the LSST alert stream
- Collaboration between NOAO and the University of Arizona (Computer Science & Mathematics Dep'ts)
- ANTARES prototype system complete and deployed
- Next step: implement public brokering of Zwicky Transient Facility public alerts (2018 or 2019)
- Supplemental funding requested from NSF to build out ANTARES system to be ready for full-scale operation in time for LSST (request pending)
- *More details in ANTARES presentation by Tom Matheson tomorrow*



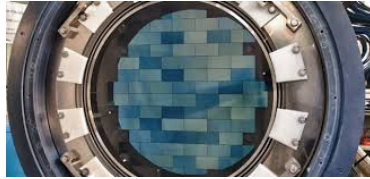


How Will Astronomers Use ANTARES?



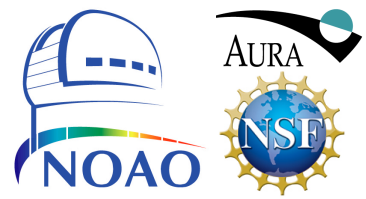


Data Management Operations (DMO)

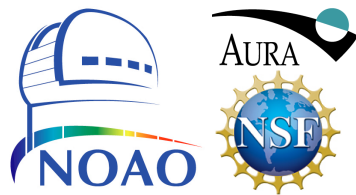


- DMO supports core transfer, archiving, and serving of data through the NOAO Science Data Archive
- Reliability improvements through ongoing modernization of data ingest and verification systems
- DECam Community Pipeline continues to deliver high-quality reduced images to PIs on short turn-around
- NOAO/CSDC exploring automated reprocessing of archival DECam data in collaboration with LBNL and NCSA (two separate efforts)
- Recent hire of Lead Web Application Developer: planning modernization of Science Data Archive interface with increased focus on services/APIs for flexibility of architecture, integration with Data Lab, and programmatic access methods for archive users

See DMO presentation by Sean McManus tomorrow for more details



OIR System Access & Coordination



Time Allocation Process

Gateway to open-access OIR observing time

- Continuing access to CTIO, KPNO, and Gemini
 - Mayall 4m transitioning to DESI installation in 2017
- Now available through NSF funding:
 - Las Cumbres Observatory Network time (via MSIP program)
 - CHARA Interferometric Array time (via MSIP program)
 - Large Binocular Telescope time (via TSIP program)
- Ongoing efforts to quantify & minimize unconscious bias
- CSDC & CTIO planning request for community input in 2017 to determine future allocation strategy for DECam after conclusion of Dark Energy Survey
- Supplemental funding requested from NSF to conduct market study for Telescope Time Exchange

See TAC presentation by Verne Smith later today





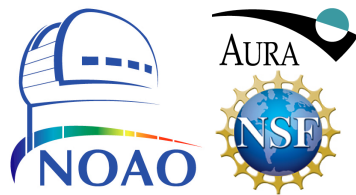
US National Gemini Office

US gateway to Gemini Observatory

- Advocates for US community interests in Gemini Observatory
- Maintains communication & coordination with Gemini staff
- Provides technical reviews and International TAC representation
- Provides helpdesk & cookbook support for Gemini data reduction
- Maintains US NGO website to publicize Gemini news & information
 - <http://ast.noao.edu/csdc/usngo>
- Organizes topical mini-workshops at January AAS meetings
 - January 2017: Mining Observatory Archives
- Currently supporting Phoenix instrument at Gemini-South

*See subsequent presentation by
Letizia Stanghellini*



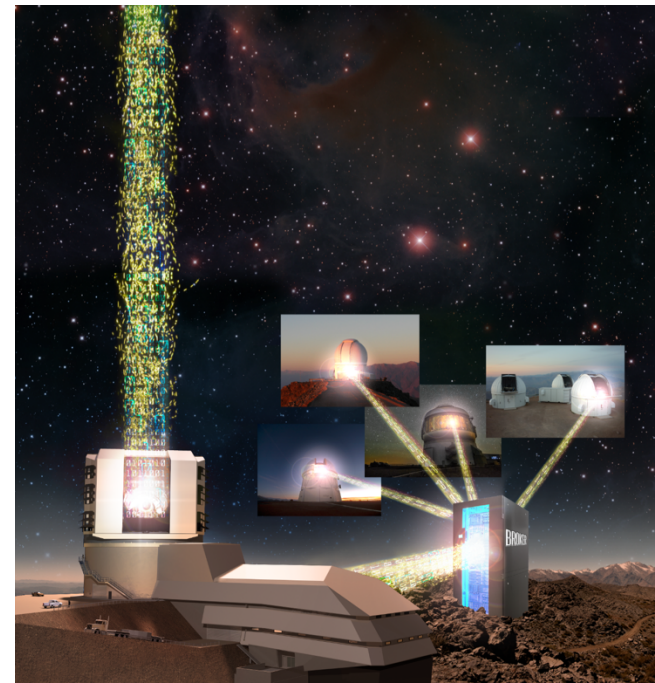


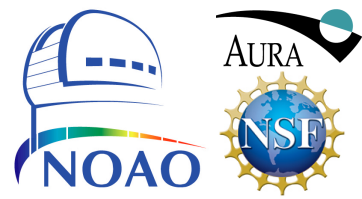
Time-Domain Infrastructure

Preparing the US OIR system for LSST follow-up

- Collaborative effort between NOAO, Las Cumbres Observatory, Gemini, SOAR, CTIO
- Integrating ANTARES broker in broader time-domain ecosystem
- Near-term time domain science opportunities with LCO, ZTF
- Hosting workshop on “Building Time Domain Infrastructure” in Tucson on 22-25 May 2017
 - <https://www.noao.edu/meetings/lsst-tds/>
- Supplemental funding requested from NSF for key elements of follow-up system:
 - Standardized communication protocols for telescope configuration & scheduling
 - Online platforms for collaborative target & observation management
 - Essential data reduction software tools

See presentation by Bob Blum later today





Other OIR System Activities

- **Decadal Survey Planning**

- NOAO effort to facilitate, coordinate, and strengthen OIR community input on 2020 decadal science and resource priorities, led by NOAO Chief Scientist Joan Najita
- Invites input on major survey projects with NOAO facilities, community access to non-NOAO facilities, access to major archival data sets, and essential resources for data-intensive and time-domain astronomy
- “Dear Colleague” letter circulated via NOAO Currents:
<https://www.noao.edu/currents/201704.html>

- **Thirty Meter Telescope Liaison Office**

- Led my Mark Dickinson of NOAO
- Draft US TMT participation plan has been submitted to NSF
- On hold pending outcome of TMT site selection



Community Development Initiatives



Education and Training



- NOAO hosts LSST Corporation-funded Data Science Fellows Program jointly with University of Arizona in April 2017
- NOAO & AURA organizing NSF-funded La Serena School for Data Science in Chile in August 2017
- Supplemental funding requested from NSF for ongoing community training in tools & methods for data-intensive astronomy at NOAO



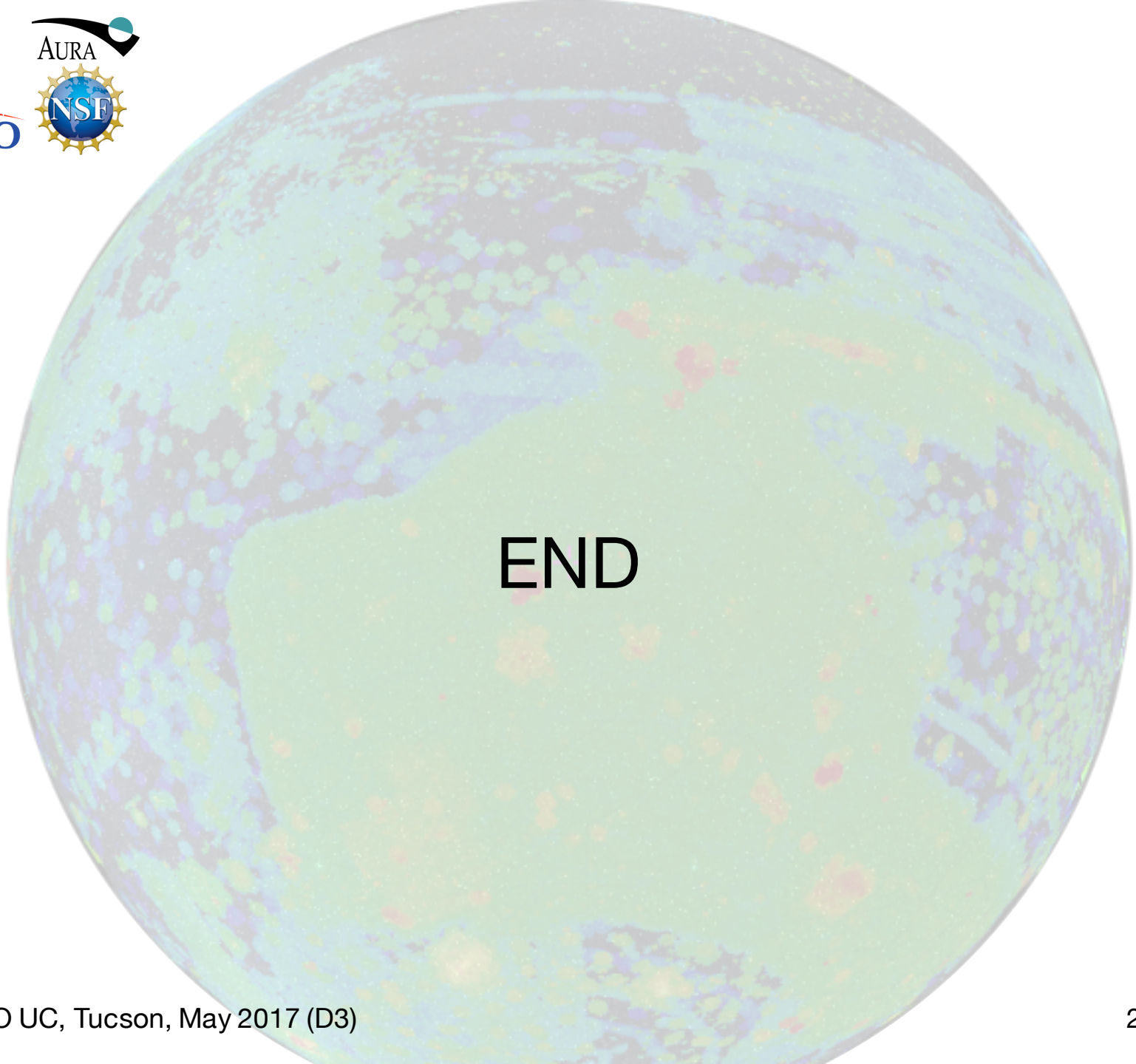
LSST Community Science Center

Development and planning

- NSF has directed NOAO to take the lead in preparing the community for LSST by developing the concept and plans for an LSST Community Science Center (LCSC)
- NOAO has recruited a diverse community-based working group to provide priorities and guidance
- The LCSC is expected to build upon the data-intensive astronomy and OIR system coordination activities currently underway at NOAO

*See (other) presentation
by Dara Norman today*





END



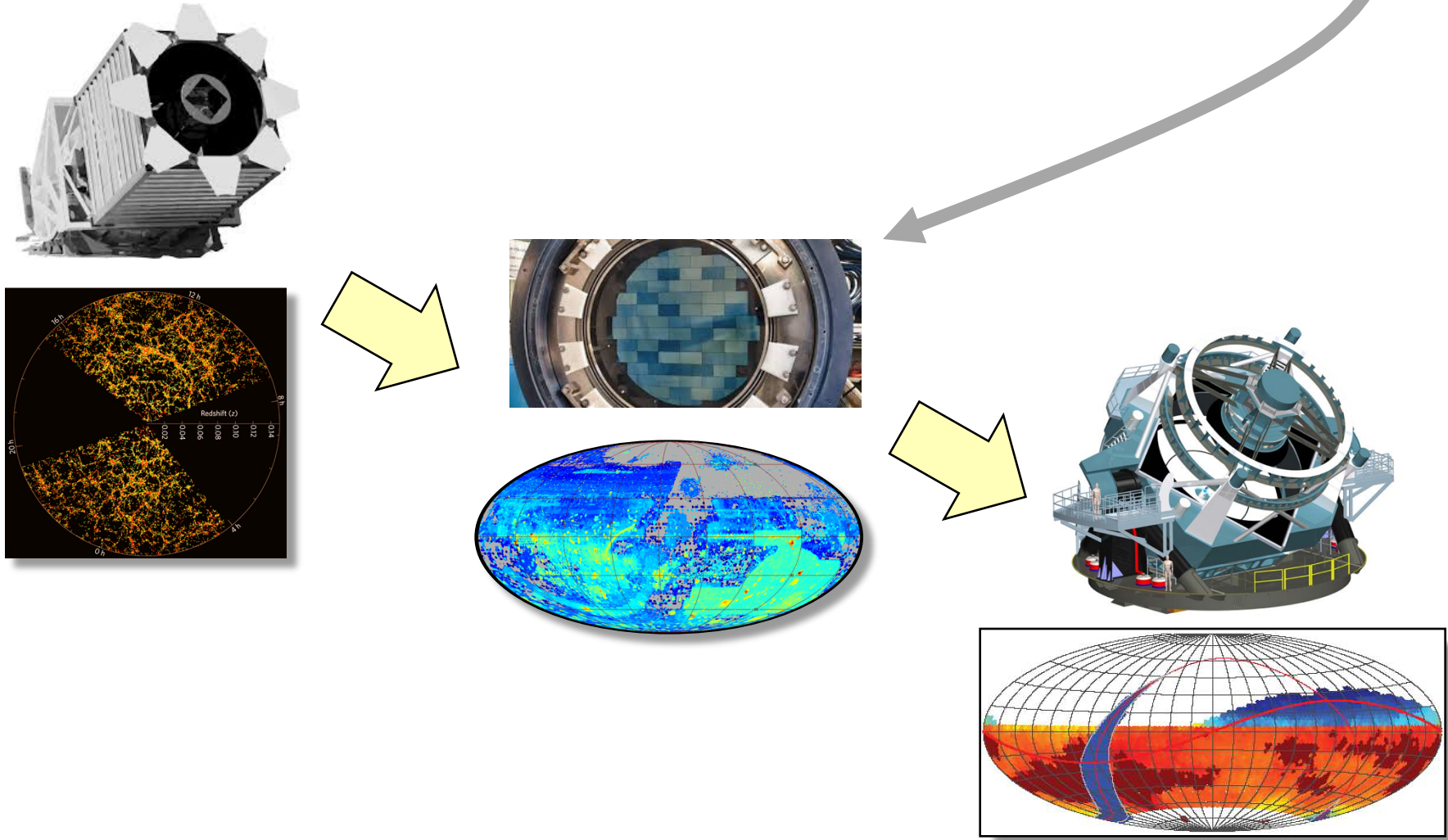
Extra Slides

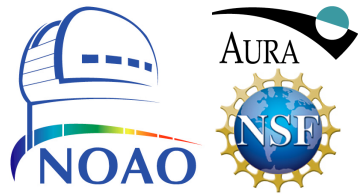


CSDC Program Overview

Supporting open science across the OIR system

- [Time Allocation Committee](#) program supports proposals for open-access observing time on telescopes across the US ground-based O/IR system through operation of a peer review process.
- [US National Gemini Office](#) serves as the interface for the US astronomical community to the Gemini Observatory.
- [Time Domain Services](#) develops infrastructure for time-domain astronomy including the ANTARES event broker.
- [Data Management Operations](#) provides core transport, archiving, processing, and online distribution for data obtained at NOAO telescopes.
- [Data Lab](#) is developing high-level tools for discovery, exploration, and analysis of large public survey datasets at NOAO.
- [Community Development](#) program encompasses the US TMT Liaison Office, the La Serena School of Data Science, and other capacity-building initiatives in US O/IR astronomy.

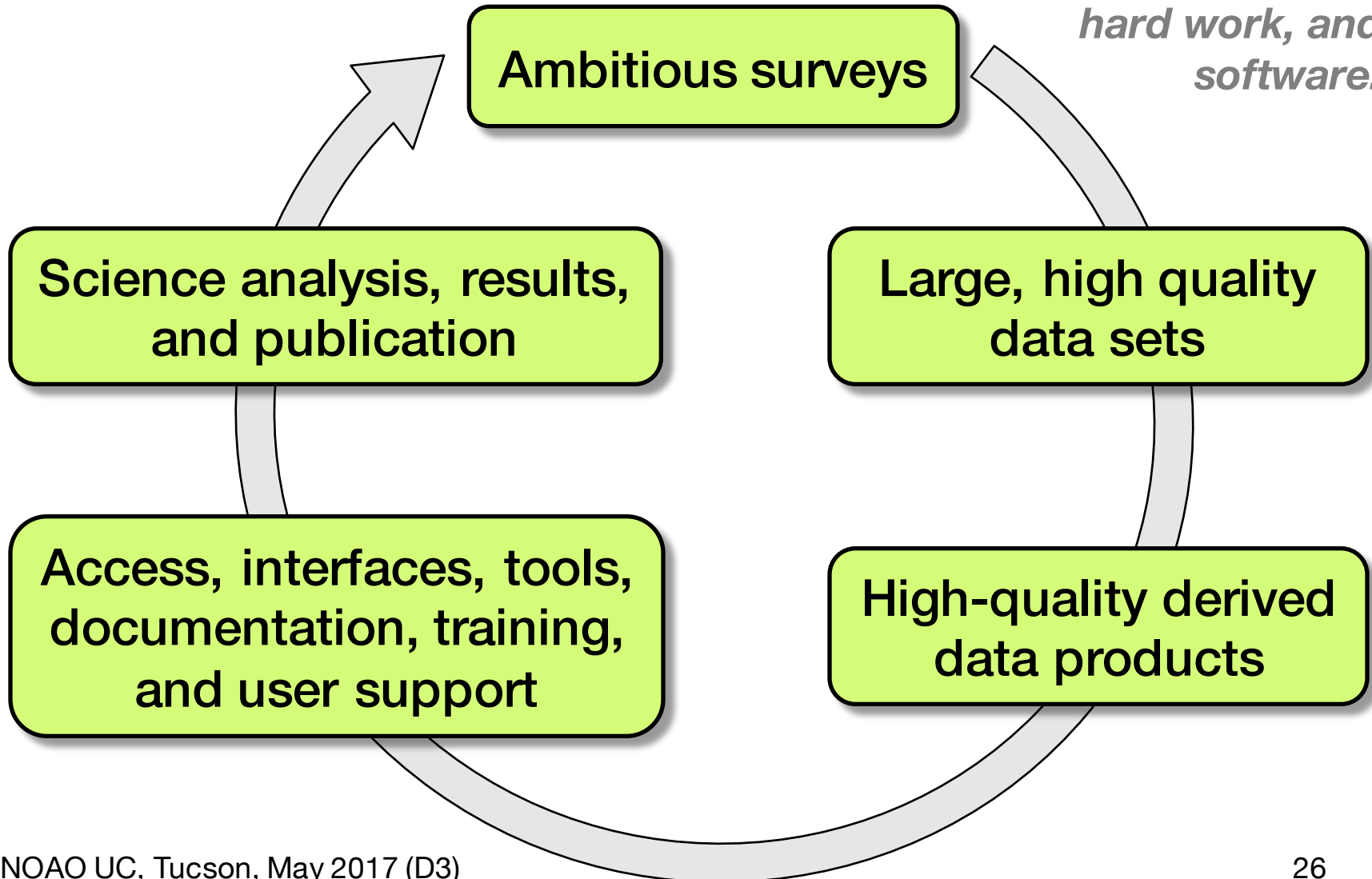


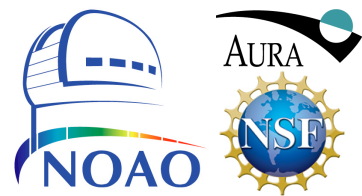


Cycle of Progress

Does not operate without all elements

*Powered by planning,
hard work, and
software!*





OK, yes, but what is Data Lab?

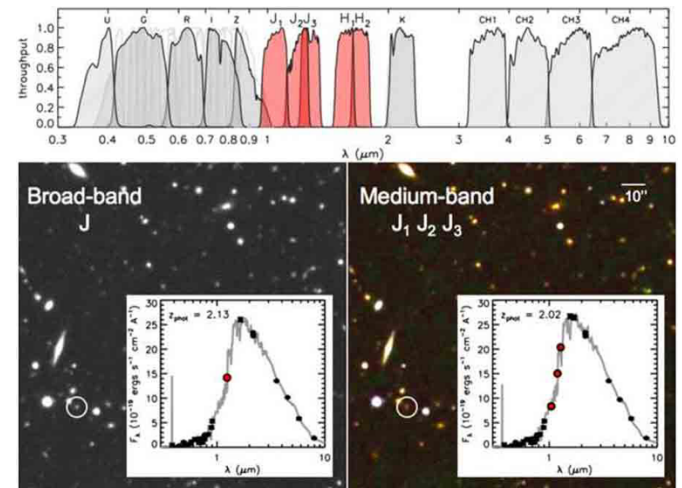
```
In [2]: # query DB for a list of all SMASH fields
query = "SELECT ra,dec,nexp,nobj FROM smash_dr1.field WHERE nexp IS NOT NUL
result = queryClient.query(token,adql=query,fmt='csv') # submit the query,
print "Inspect the returned result: it's a csv-formatted string:"
print result[:105]
```

- Baseline: provide SDSS-like interfaces to NOAO survey datasets
- *But not just that:* Data Lab part of a broader “[Archive 2.0](#)” trend:
 - [Big data is the rule](#): support for *all* users—not just power users—to work with large databases and data archives
 - [Code-to-data frameworks](#): containerization, Jupyter notebooks, server-side virtual storage, cloud-like provisioning models
 - [Archive-as-platform](#): standardized services and application programming interfaces (APIs) enable flexibility, integration with diverse applications, and crowdsourcing of development
- This trend is distributed across astronomy (including within LSST).
- Think of this technology in astronomy today like personal computers ca. 1980 or smartphones ca. 2005 in the wider world.
- We have the data, we have the people, we have the mission.
NOAO must [share and collaborate](#); NOAO must also [lead](#).

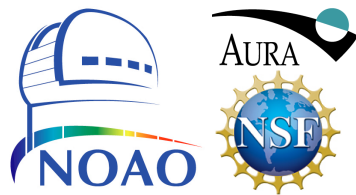
Major Survey Datasets at NOAO

Other NOAO survey programs

- Currently collecting, curating, and migrating historical NOAO survey data archives to new platform
- Surveys under curation for updated release in 2017:
 - ChaMPlane, Deep Ecliptic Survey, Deeperange Survey, Deep Lens Survey, Faint Sky Variability Survey, FLAMEX, Infrared Bootes Imaging Survey, NEWFIRM Medium Band Survey, NOAO Deep Wide Field Survey, NOAO Fundamental Plane Survey, Z Bootes Imaging Survey
- Additional surveys forthcoming
- NOAO aims to provide long-term archival and research-enabling hosting of diverse ground-based OIR survey data sets
- Effort led by Tod Lauer of NOAO



Whitaker et al. (2011) / NMBS



Partnerships and Connections

We can't do this alone!

LSST	AD-CSDC on LSSTC Exec Board; AD-CSDC visit to UW in Sep 2016; LSST DM rep on ANTARES review; community-science discussions with LSST Deputy Director; LSST rep to be on NOAO LSST Community Science WG; anticipating systems & services convergence in future; next steps await finalization of LSST DM re-plan and onboarding of new LSST DM head.
STScI	AD-CSDC visit to STScI in April 2017, coordinating with new STScI Data Science Mission Head; Data Lab Program Head on MAST Users' Group; NOAO participation in STScI Feb. 2017 Big Data meeting; STScI Python community software initiative interest coordination.
Pan-AURA	Participation in pan-AURA working groups on (i) virtualization & cyber-infrastructure and (ii) Python in astronomy.
DES/NCSA	Collaboration between DES-DM and NOAO DECam pipelines; negotiation of DES public-release terms; partnership role in DES DR1 in Dec 2017; multiple bilateral staff visits NOAO-NCSA.
DESI/NERSC	DESI targeting survey coordination with Data Lab for database support and data releases; NOAO pipeline processing of DESI target imaging data; data-transfer coordination with NERSC; planning to mirror DESI data at NOAO; planning to deploy high-level DESI data interfaces based on SDSS.
SDSS	Mirroring public SDSS data at NOAO; building SDSS hooks into Data Lab; evaluating SDSS interfaces for adaptation to NOAO; collaboration with SDSS archive sustainability and transferability initiative; working towards NOAO role as long-term home of SDSS archive.
IVOA/ADASS	Regular ADASS attendance by CSDC staff; evaluating and adopting VO protocols where appropriate; AD-CSDC to be US ground-based representative to IVOA Executive.
U. Of Arizona	Collaboration with CS and Math in ANTARES project; using UITS co-location facility for archive mirror and ANTARES system; hosting LSST Data Science Fellows jointly with Steward Obs.
Gemini, SOAR, Las Cumbres	Collaboration on LSST follow-up infrastructure planning and development (<i>Blum talk</i>); exploring avenues for software-development sharing and coordination; planning time-allocation evolution.



User engagement

FY17 highlights

- Data Lab presence and demonstrations at AAS meetings in January and June 2017
- US NGO AAS session on archival research opportunities (Jan 2017)
- Planning pre-LSST ANTARES engagement with Zwicky Transient Facility and science users of its community alert stream
- LSST Community Science Center working group forming
- Supporting Data Lab in-house “early adopter” programs
- Data Lab feature at STScI Big Data meeting in Feb/Mar 2017
- Data Lab and ANTARES modules for LSST Data Science Fellows Program at NOAO in April 2017
- Data Lab pre-release tutorial sessions in May 2017
- Engagement with DESI collaboration high-level data distribution use cases (imaging and spectroscopy)
- Web design and online documentation development
- General AD-CSDC evangelism, jawboning, and conversation