

## The NOAO Science Archive

Mark Dickinson for NOAO Science Data Management



### NSA ≠ NSA

## NOAO Science Archive







### NSA ≠ NSA

# NOAO Science Archive







#### **NOAO Science Archive**

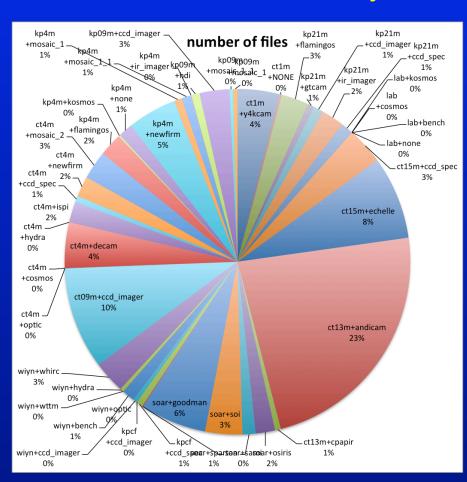
- Raw data from:
  - 11 telescopes (KPNO, CTIO, SOAR, WIYN)
  - 48 telescope+instrument combinations
- Data from 2004B to the present
  - Earlier raw data are stored on tape
- Three geographically dispersed copies of raw data
- Pipeline-reduced data from NOAO wide-field imagers:
  - Mosaic (1, 1.1, 2, at KPNO 4m+0.9m and CTIO 4m)
  - NEWFIRM (KPNO 4m + CTIO 4m)
  - DECam (CTIO 4m)
- Delivered, reduced survey data products
  - Older survey data currently in separate (older) Survey Archive
  - New DECam survey data (including DES) go into main Science Archive

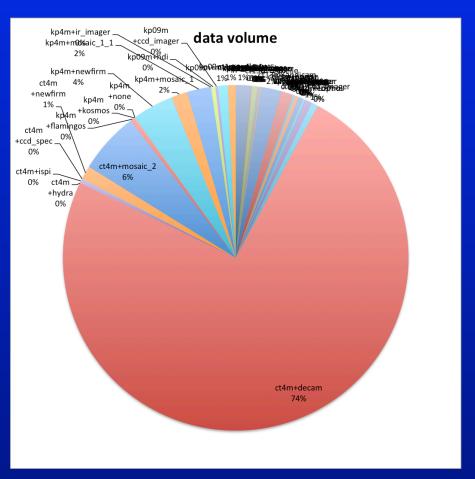


#### NOAO Science Archive - Data Holdings

6,661,685 raw files (184.7 TB, compressed) 2,773,750 reduced files (149.6 TB, compressed)

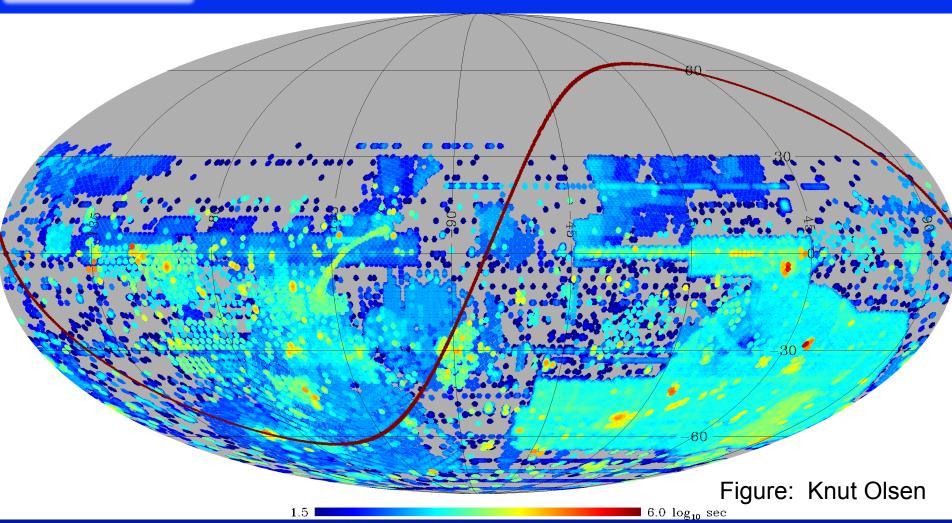
#### Raw data by telescope + instrument:





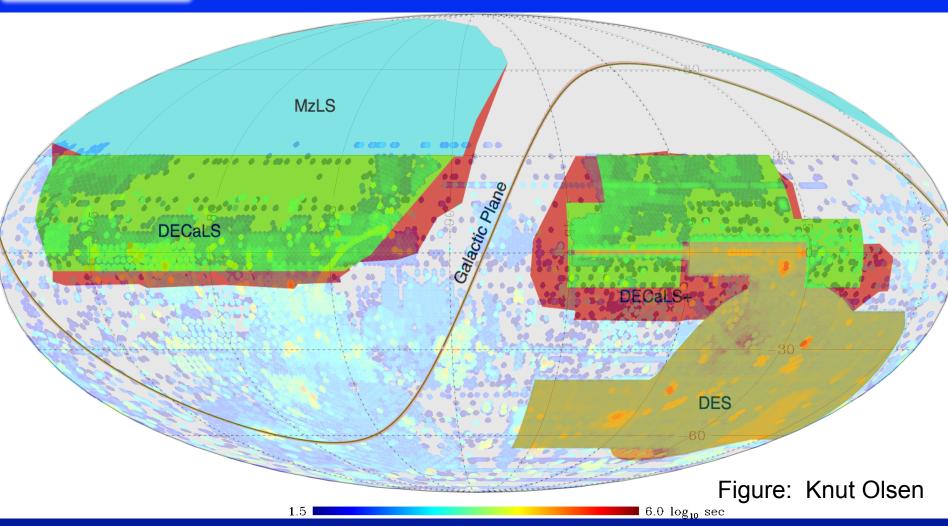


### Current DECam data holdings





### Incoming wide-field survey data





Rapidly growing hoard of NOAO wide-field imaging is a rich trove.

NOAO Archive is transitioning from being mainly a storage & distribution service (PI-oriented) to a data resource (archival users).

#### Mining the NOAO Archive

arXiv.org > astro-ph > arXiv:1503.02079

Search or A

Astrophysics > Astrophysics of Galaxies

#### Beasts of the Southern Wild. Discovery of a large number of Ultra Faint satellites in the vicinity of the Magellanic Clouds

Sergey E. Koposov, Vasily Belokurov, Gabriel Torrealba, N. Wyn Evans

(Submitted on 6 Mar 2015 (v1), last revised 10 Mar 2015 (this version, v2))

We have used the publicly released Dark Energy Survey data to hunt for new satellites of the Milky Way in the Southern hemisphere. Our search yielded a large number of promising candidates. In this paper, we announce the discovery of 9 new unambiguous ultra-faint objects, whose authenticity can be established with the DES data alone. Based on the morphological properties, three of the new satellites are dwarf galaxies, one of which is located at the very outskirts of the Milky Way, at a distance of 380 kpc. The remaining 6 objects have sizes and luminosities comparable to the Segue 1 satellite and can not be classified straightforwardly without follow-up spectroscopic observations. The satellites we have discovered cluster around the LMC and the SMC. We show that such spatial distribution is unlikely under the assumption of isotropy, and, therefore, conclude that at least some of the new satellites must have been associated with the Magellanic Clouds in the past.

arXiv.org > astro-ph > arXiv:1503.02584

Search or

Astrophysics > Astrophysics of Galaxies

#### Eight New Milky Way Companions Discovered in First-Year Dark Energy Survey Data

The DES Collaboration, K. Bechtol, A. Drlica-Wagner, E. Balbinot, A. Pieres, J. D. Simon, B. Yanny, B. Santiago, R. H. Wechsler, J. Frieman, A. R. Walker, P. Williams, E. Rozo, E. S. Rykoff, A. Queiroz, E. Luque, A. Benoit-Levy, R. A. Bernstein, D. Tucker, I. Sevilla, R. A. Gruendl, L. N. da Costa, A. Fausti Neto, M. A. G. Maia, T. Abbott, S. Allam, R. Armstrong, A. H. Bauer, G. M. Bernstein, E. Bertin, D. Brooks, E. Buckley-Geer, D. L. Burke, A. Carnero Rosell, F. J. Castander, C. B. D'Andrea, D. L. DePoy, S. Desai, H. T. Diehl, T. F. Eifler, J. Estrada, A. E. Evrard, E. Fernandez, D. A. Finley, B. Flaugher, E. Gaztanaga, D. Gerdes, L. Girardi, M. Gladders, D. Gruen, G. Gutierrez, J. Hao, K. Honscheid, B. Jain, D. James, S. Kent, R. Kron, K. Kuehn, N. Kuropatkin, O. Lahav, T. S. Li, et al. (32 additional authors not shown)

(Submitted on 9 Mar 2015)

We report the discovery of eight new Milky Way companions in ~1,800 deg^2 of optical imaging data collected during the first year of the Dark Energy Survey (DES). Each system is identified as a statistically significant over-density of individual stars consistent with the expected isochrone and luminosity function of an old and metal-poor stellar population. The objects span a wide range of absolute magnitudes (M\_V from -2.2 mag to -7.4 mag), physical sizes (10 pc to 170 pc), and heliocentric distances (30 kpc to 330 kpc). Based on the low surface brightnesses, large physical sizes, and/or large Galactocentric distances of these objects, several are likely to be new ultra-faint satellite galaxies of the Milky Way and/or Magellanic Clouds. We introduce a likelihood-based algorithm to search for and characterize stellar overdensities, as well as identify stars with high satellite membership probabilities. We also present completeness estimates for detecting ultra-faint galaxies of varying luminosities, sizes, and heliocentric distances in the first-year DES data.



# Raw DECam data in Archive (as of today)

- 273,886 raw DECam files (139.5 TB)
  - DECam already makes up 75% of the total raw data volume in the Archive
  - As many DECam exposures taken since December 2012 as with the CTIO+KPNO Mosaic cameras combined since 2004B !!
  - 192,258 public (as of today) (70%)
  - 173,446 "object" exposures (62%; the rest are calibration, etc.)
  - 75,931 exposures come from DES (28%)
    - 51,816 DES frames are public (68% of total DES)



# Pipeline-reduced DECam data in Archive (as of today)

- 782,929 reduced DECam data files
  - Note: CP produces several reduced files for each raw exposure
  - 387,456 reduced public (49%)
- Reduced single-frame "object" images:
   (PROCTYPE = InstCal, PRODTYPE = image)
  - 129,188 exposures
    - 12,027 are delivered products from DES-Y1 (9%; all public)
    - 117,161 from CP (90%)
      - 58,253 CP public (50%)
- Reduced coadds:

```
(PROCTYPE = Stacked, PRODTYPE = image)
```

- 16,415 stacks
  - 8,304 public (51%)



# NOAO Science Archive http://portal-nvo.noao.edu



#### NOAO Science Archive

Home Contents

Search NOAO data

Contributed software

Help

#### Home »

#### Manage your user account | Log in

#### Latest News

Feburary 2015

Dark Energy Survey (DES) reduced data products available.

Reduced data from the Dark Energy Survey (DES) Year 1 public data release are now available in the NOAO Science Archive. The data release consists of a large number of DECam images in several sky areas, individually reduced and calibrated (photometrically and astrometrically) by the DES science team. The Simple Query Form now includes buttons that simplify searches for both raw and reduced DES images. Observations have a 12 month proprietary period; by February 2015, all data will be public. See the DES Year 1 data release notes for full documentation.

#### September 2013

#### New username and login system.

The NOAO Archive has a new login and security system, which provides easier and quicker sign-on for accessing proprietary data. Pls and co-ls of NOAO observing programs will automatically be assigned new archive usernames and passwords. If you have previously used an NVO username to sign in to the Archive, you may continue to do so, or switch to your new NOAO Archive username. You do not need to log in to query or retrieve public, non-proprietary data.

The NOAO Science Archive provides access to data from more than 25 NOAO telescope + instrument combinations, including those operated in partnership with the WIYN, SOAR, and SMARTS consortia. Pipeline-reduced data products from the DECam, Mosaic and NEWFIRM imagers on the KPNO and CTIO 4m telescopes are available.

Proprietary data access (login required)

General search for NOAO data (all users)

Search NOAO Survey Program high-level data products



# NOAO Science Archive http://portal-nvo.noao.edu



#### NOAO Science Archive

Home Contents

Search NOAO data

Contributed software

Help

#### Home »

Latest News

Feburary 2015

Dark Energy Survey (DES) reduced data products available.

Reduced data from the Dark Energy Survey (DES) Year 1 public data release are now available in the NOAO Science Archive. The data release consists of a large number of DECam images in several sky areas, individually reduced and calibrated (photometrically and astrometrically) by the DES science team. The Simple Query Form now includes buttons that simplify searches for both raw and reduced DES images. Observations have a 12 month proprietary period; by February 2015, all data will be public. See the DES Year 1 data release notes for full documentation.

September 2013

#### New username and login system.

The NOAO Archive has a new login and security system, which provides easier and quicker sign-on for accessing proprietary data. Pls and co-ls of NOAO observing programs will automatically be assigned new archive usernames and passwords. If you have previously used an NVO username to sign in to the Archive, you may continue to do so, or switch to your new NOAO Archive username. You do not need to log in to query or retrieve public, non-proprietary data.

The NOAO Science Archive provides access to data nom ...... then OF NOAO tolescope +

Manage your user account | Log in

instrument combinations, including those operated in partnership with the WIYN, SOAR, and SMARTS consortia. Pipeline-reduced data products from the DECan, Mosaic and NEWFIRM imagers on the KPNO and CTIO 4m telescopes are available.

Proprietary data access (login required)

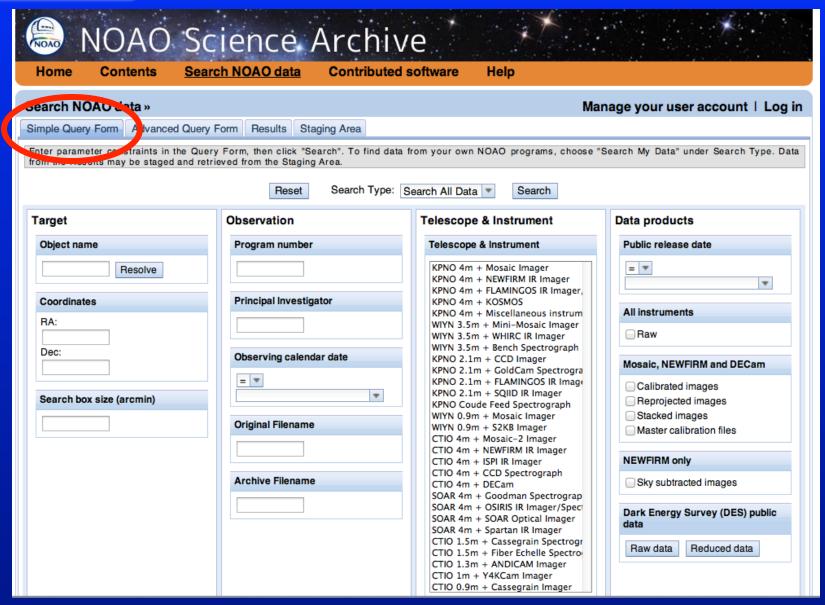
General search or NOAO data (all users)

Search NOAO Survey Program high-level data products

PIs and co-Is of NOAO programs log in to access their proprietary data



### Simple Query Form





### Simple Query Form

Heset Search Type: Search All Data V Search				
Observation	Telescope & Instrument	escope & Instrument Data products		
Program number	Telescope & Instrument	Public release date		
	KPNO 4m + Mosaic Imager KPNO 4m + NEWFIRM IR Imager KPNO 4m + FLAMINGOS IR Imager	=		
Principal Investigator	KPNO 4m + KOSMOS KPNO 4m + Miscellaneous instrum WIYN 3.5m + Mini-Mosaic Imager	All instruments  Raw		
Observing calendar date	WIYN 3.5m + WHIRC IR Imager WIYN 3.5m + Bench Spectrograph KPNO 2.1m + CCD Imager KPNO 2.1m + GoldCam Spectrogra	Mosaic, NEWFIRM and DECam		
=	KPNO 2.1m + FLAMINGOS IR Image KPNO 2.1m + SQIID IR Imager KPNO Coude Feed Spectrograph	Calibrated images Reprojected images		
Original Filename	WIYN 0.9m + Mosaic Imager WIYN 0.9m + S2KB Imager CTIO 4m + Mosaic-2 Imager CTIO 4m + NEWFIRM IR Imager	Stacked images Master calibration files		
Archive Filename	CTIO 4m + ISPI IR Imager CTIO 4m + CCD Spectrograph CTIO 4m + DECam SOAR 4m + Goodman Spectrograp	NEWFIRM only  Sky subtracted images		
	SOAR 4m + Goodman Spectrograp SOAR 4m + OSIRIS IR Imager/Spect SOAR 4m + SOAR Optical Imager SOAR 4m + Spartan IR Imager	Dark Energy Survey (DES) public data		
	CTIO 1.5m + Cassegrain Spectrogr CTIO 1.5m + Fiber Echelle Spectro CTIO 1.3m + ANDICAM Imager CTIO 1m + Y4KCam Imager CTIO 0.9m + Cassegrain Imager	Raw data Reduced data		



### Simple Query Form

Reset Search Type: Search All Data Search				
Observation	Telescope & Instrument	Data products		
Program number	Telescope & Instrument	Public release date		
DES	KPNO 4m + Mosaic Imager KPNO 4m + NEWFIRM IR Imager KPNO 4m + FLAMINGOS IR Imager,	2015-03-11		
Principal Investigator	KPNO 4m + KOSMOS KPNO 4m + Miscellaneous instrum	All instruments		
	WIYN 3.5m + Mini-Mosaic Imager			
	WIYN 3.5m + WHIRC IR Imager WIYN 3.5m + Bench Spectrograph	Raw		
Observing calendar date	KPNO 2.1m + CCD Imager KPNO 2.1m + GoldCam Spectrogra	Mosaic, NEWFIRM and DECam		
= 🔻	KPNO 2.1m + FLAMINGOS IR Image KPNO 2.1m + SQIID IR Imager KPNO Coude Feed Spectrograph	✓ Calibrated images  ☐ Reprojected images ☐ Stacked images ☐ Master calibration files		
Original Filename	WIYN 0.9m + Mosaic Imager WIYN 0.9m + S2KB Imager CTIO 4m + Mosaic-2 Imager			
	CTIO 4m + NEWFIRM IR Imager CTIO 4m + ISPI IR Imager	NEWFIRM only		
Archive Filename	CTIO 4m + CCD Spectrograph CTIO 4m + DECam SOAR 4m + Goodman Spectrograp	Sky subtracted images		
	SOAR 4m + OSIRIS IR Imager/Spect SOAR 4m + SOAR Optical Imager SOAR 4m + Spartan IR Imager	Dark Energy Survey (DES) public data		
	CTIO 1.5m + Cassegrain Spectrogr CTIO 1.5m + Fiber Echelle Spectro CTIO 1.3m + ANDICAM Imager CTIO 1m + Y4KCam Imager CTIO 0.9m + Cassegrain Imager	Raw data Reduced data		



#### **Advanced Query Form**



### NOAO Science Archive

Home

Contents

Search NOAO data

Contributed software

Help

#### Search NOAO data

Manage your user account | Log in

Simple Query Form

Advanced Query Form

Pesults

Staging Area

Enter your SQL into in Query Text Area, the click "Search". Data from the Results may be staged and retrieved from the Staging Area.

You may enter parameters into the Simple Query Form, then switch to the Advanced Query Form to see the equivalent SQL. This can provide useful examples that you can modify to form your own custom queries.

There are a few required columns that you must include in your SELECT clause, these are: reference, release\_date, start\_date, filesize, dtpropid, and md5sum. Only voi.siap and voi.skyimage tables are allowed.

For further information about the schema of the tables voi.siap and voi.skyimage, please click here. Also, if you want to know more about SQL please click here or here for advanced users.

Warning: Queries by default will return data for which you are PI, and/or public data. To obtain data for which you are co-I, see the SQL example below. Observing mode is only applicable to KOSMOS instruments. The Advanced Query functionality is a new feature and will be improved in future versions of the archive.

#### Error trying to send your query.

Reset

Search

SELECT reference, dtpropid, surveyid, release\_date, start\_date, date\_obs, dtpi, ra, dec, telescope, instrument, filter, exposure, obstype, obsmode, proctype, prodtype, seeing, depth, dtacqnam, reference AS archive\_file, filesize, md5sum FROM voi.siap WHERE (release\_date <= '2015-03-11' AND (proctype = 'InstCal') AND (prodtype IS NULL OR prodtype <> 'png')) AND (dtpropid ILIKE '%DES%' OR surveyid ILIKE '%DES%') ORDER BY date\_obs ASC LIMIT 20000

#### Examples of valid SQL queries

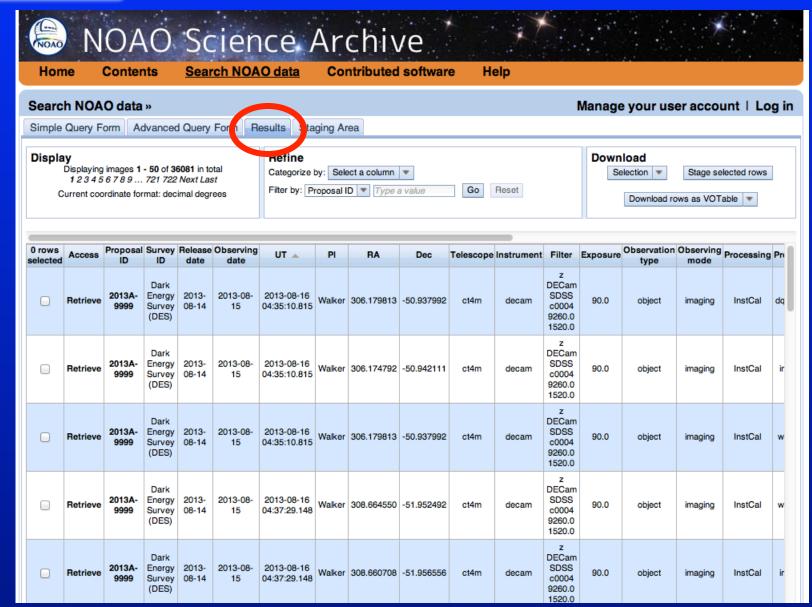
 This selects the first 100 rows from the voi.siap table within a box bounded by ra (10.352, 11.017) and dec (41.019, 41.519). Note that position constraints must be in the same units (here, degrees). Tip: LIMIT 100 or some other small number is useful for testing without getting a very large data set back.

SELECT \* FROM voi.siap WHERE ((ra >= 10.352 AND ra <= 11.017) AND (dec >= 41.019 AND dec <= 41.519)) LIMIT 100

This selects the first 100 rows from the voi.siap table for a specified NOAO proposal ID (here, noao) using a wildcard match. A wildcard may be specified by constructing a condition using the operator "LIKE" or "ILIKE" (the later is case insensitive) followed by a string which may contain one or more "%" characters which are the wildcards.
 SELECT \* FROM voi.siap WHERE dtpropid ILIKE '%noao%' LIMIT 100

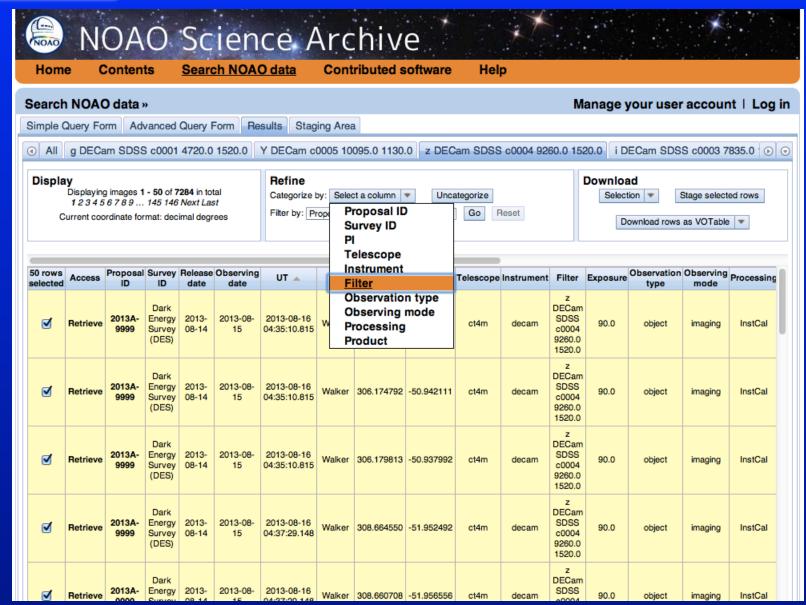


#### Search Results





#### Search Results





#### **Staging Area**

Manage your user account | Log in



Displaying images 1 - 20 of 60 in total

Launch download manager

Help

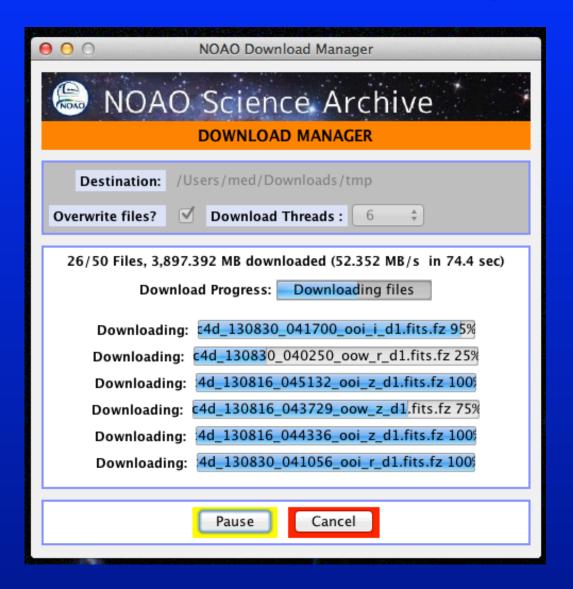
File name	File size	Stage status
c4d_130816_043510_oow_z_d1.fits.fz	162.608 [MB]	Staged
c4d_130816_043510_ooi_z_d1.fits.fz	318.453 [MB]	Staged
c4d_130816_043510_ood_z_d1.fits.fz	9.521 [MB]	Staged
c4d_130816_043729_ooi_z_d1.fits.fz	318.246 [MB]	Staged
c4d_130816_043729_oow_z_d1.fits.fz	162.181 [MB]	Staged
c4d_130816_043729_ood_z_d1.fits.fz	8.772 [MB]	Staged
c4d_130816_043930_ood_z_d1.fits.fz	8.787 [MB]	Staged
c4d_130816_043930_ooi_z_d1.fits.fz	318.067 [MB]	Staged
c4d_130816_043930_oow_z_d1.fits.fz	161.980 [MB]	Staged
c4d_130816_044136_ood_z_d1.fits.fz	8.289 [MB]	Staged
c4d_130816_044136_ooi_z_d1.fits.fz	317.716 [MB]	Staged
c4d_130816_044136_oow_z_d1.fits.fz	161.839 [MB]	Staged
c4d_130816_044336_oow_z_d1.fits.fz	162.475 [MB]	Staged
c4d_130816_044336_ood_z_d1.fits.fz	8.925 [MB]	Staged
c4d_130816_044336_ooi_z_d1.fits.fz	317.650 [MB]	Staged
c4d_130816_044541_ood_z_d1.fits.fz	9.046 [MB]	Staged
c4d_130816_044541_oow_z_d1.fits.fz	161.663 [MB]	Staged
c4d_130816_044541_ooi_z_d1.fits.fz	317.929 [MB]	Staged
c4d_130816_044831_ooi_z_d1.fits.fz	318.905 [MB]	Staged
c4d_130816_044831_oow_z_d1.fits.fz	162.547 [MB]	Staged



#### **Download Manager**

- Fast parallel transfer
- Easy restart if needed

Next release will automatically pull files from mass storage in North or South depending on user location.





#### Help and Documentation



#### NOAO Science Archive: Tutorials:

Return To Portal Search and Retrieve NOAO Data PI & Co-I Data Access Data compression Archive file names

#### Search and Retrieve NOAO Data

The *Query page* lets you search for data in the NOAO Science Archive. The Archive holds data from many different combinations of telescopes and instruments, including the NOAO facilities at **KPNO** and **CTIO**, and NOAO data from consortium facilities such as **WIYN**, **SOAR** and **SMARTS**. All raw data from these telescopes and instruments are archived, as well as pipeline-reduced data products from the DECam, Mosaic and NEWFIRM imagers on the NOAO 4m telescopes. Principal Investigators and authorized co-investigators of NOAO observing programs who have registered with the NOAO Archive can retrieve their proprietary data using the Query form. Any user (registered or not) can search for and retrieve non-proprietary data as well.

The process of finding and accessing NOAO data follows several basic steps:

- . Searching for data using the query form
  - Search for data from your own NOAO observing programs
  - Search for any data in the NOAO Archive
  - Search form parameters
- Search results
  - Sorting, filtering, and categorizing search results
  - Selecting data for retrieval
- Staging data for ftp retrieval
  - The Staging Area
  - Retrieving your data
  - Problems with staging data
  - Cleaning up your ftp staging area
- Retrieving data directly from the Archive using cURL
- · Working with your downloaded data
  - Data compression
  - Archive filenames and renaming

Also, send email to: sdmhelp@noao.edu

#### The Search Form



 At top left (under Search NOAO data), there are four tabs marked Query Form, Advanced Query Form, Results, and Staging Area. These let you switch back and forth between the search form (or the advanced query form), the search results, and the place where you will stage selected data for ftp retrieval.



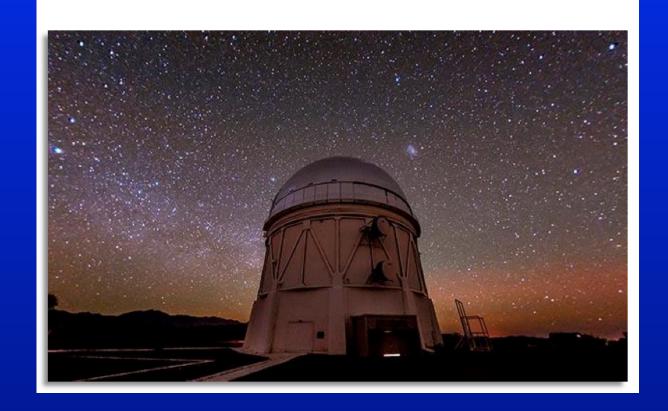
## NOAO Data Handbook

#### NOAO DATA HANDBOOK

VERSION 2.0, FEBRUARY 2015

#### Newly revised! R. Shaw, editor

- Intro to NOAO data and archives
- Mosaic cameras
- NEWFIRM
- DECam





### **Archive Usage**

- Download statistics for CY2014 (approximate):
  - 717,000 files downloaded
  - 76 TB downloaded
  - average of 150 FTP users per month



#### What's next for the NOAO Archive?

- Features in upcoming archive releases:
  - Larger staging volumes (up to 250,000 files)
  - Better performance
- New survey data:
  - DECam Legacy Survey (DECaLS) DR1 this month!
  - Other DECam data from NOAO Survey Programs
  - Eventually: migrate pre-DECam surveys from old (obsolete)
     Survey Archive to NOAO Science Archive
- Behind the scenes:
  - Major overhaul of 3<sup>rd</sup>-party software systems & infrastructure
- On my long-term wish-list:
  - Uniform reprocessing of all Mosaic, NEWFIRM and DECam data



### Big images lead to big catalogs

- NOAO imaging data is growing to ~petabyte scale:
  - Dark Energy Survey
  - DECaLS and DESI Targeting Survey
  - Community DECam programs and surveys
- NOAO Science Archive serves images, not catalogs
- But...Large catalogs are coming:
  - Dark Energy Survey 45 TB
  - DESI Targeting Survey − ~5 TB
  - Community programs and surveys up to several TB each



#### The NOAO Data Lab

# Enable efficient exploration and analysis of the large datasets being generated by instruments on NOAO widefield 4-m telescopes:

- Science with large catalogs
- Easy interplay between catalog objects & associated image data
- Combine services into custom workflows to derive scientific results from catalogs + pixels
- Collaboration tools for distributed research teams working on large data sets

#### The Data Lab team (so far):

- Project scientist: Knut Olsen
- System architect & lead developer: Mike Fitzpatrick
- In the trenches: Matthew Graham, Ken Mighell, Betty Stobie, Pat Norris, Stephen Ridgway



#### The Data Lab in a Nutshell

Large Catalogs – Data Lab will serve TB-scale databases

Pixel Data – Data Lab will connect users to images and spectra in NOAO Science Archive

Virtual Storage – Minimizes data transfer

Visualization – Data Lab will enable data exploration

Compute Processing – Data Lab will allow workflows to run close to the data

Additional features – Access to published datasets and external data services, data publication, exportable workflows, distributable software



#### **NOAO Data Lab**

- Data Lab will benefit from and build off existing tools and systems where possible
  - SDSS CasJobs, MAST, IRSA, CANFAR, NERSC, VAO, community-led projects

#### Timeline:

- Conceptual design review next week!
- Science demonstration at AAS January 2016
  - Functional prototypes of various subsystems
- First public release in January 2017



#### **Discussion session**

Friday, 10:45 – 11:45

Round-table discussion 7b: NOAO Science Archive

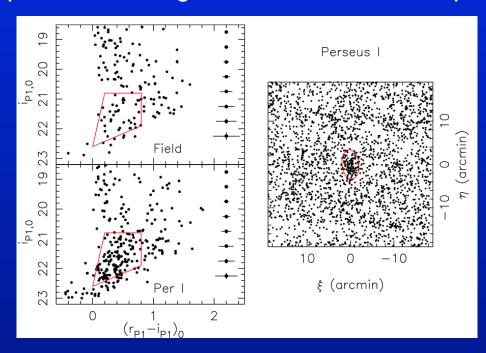
- What is good?
- What should the NOAO Archive be doing better?
- What isn't the NOAO Archive doing at all?



#### Extra slides



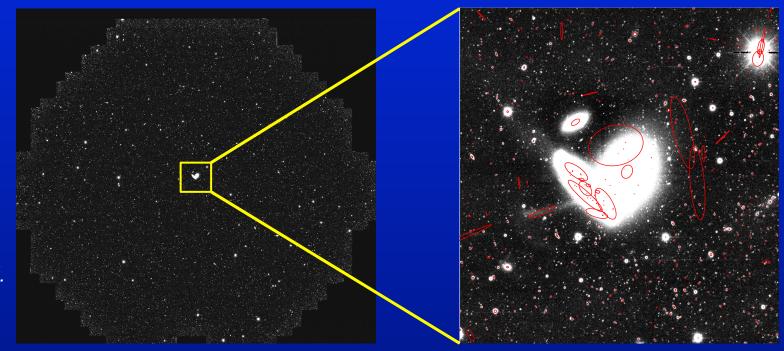
- Catalog science
  - Example: search for Galactic substructure through photometric selection of candidate populations
  - Data Lab will provide access to large catalog databases, query interface, personal storage, and visualization capability





#### Exploration

- Example: selection of a sample of large galaxies from a catalog, retrieving image cutouts, overlaying with catalog measurements
- Data Lab will provide a fast image cutout service, visualization capability, cross-match service



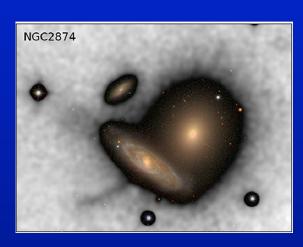
NGC 2874 Sweet et al. DECam program

DECam Community Workshop - 11 March 2015

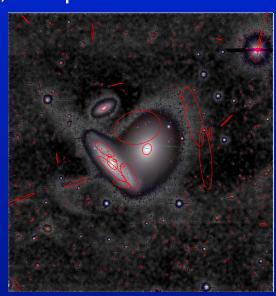


#### Custom workflows

- Example: Use a large sample of galaxies to determine frequency of minor mergers, obtaining image cutouts and performing custom pixel analysis (e.g. PSF subtraction, image filtering, automated feature detection)
- Data Lab will provide ability to string previous services into a workflow, interface to legacy software, compute service



Miskolczi et al. (2011)



NGC 2874 Sweet et al. DECam program

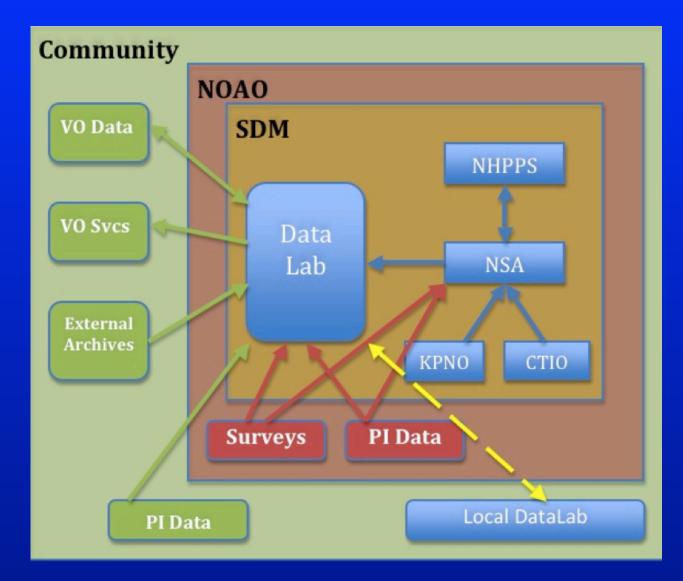


#### Collaborative research

- Example: SMASH collaboration has 30 investigators working on many aspects of search for Magellanic Cloud populations all over the sky. Data products include reduced images and photometry from three different approaches; tasks include star/galaxy separation, foreground modeling, population detection, population analysis, and simulations, divided over different people
- Data Lab will provide access to a shared storage space, shareable containers for software tasks, and a customizable virtual machine

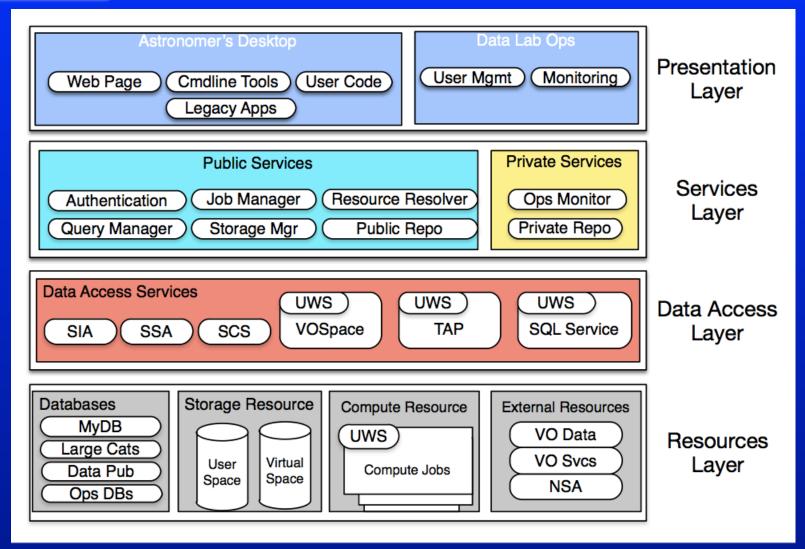


#### **Data Lab Context**





#### Data Lab Architecture





#### Data Lab System Architecture

Example Flow of Use Case through System Architecture

