US National Gemini Office (NGO) Program Update

Letizia Stanghellini
NGO

UC Meeting, Tucson, May 16-17, 2017

NGO
mission

- The US NGO supports Gemini users in the various phases of science program planning and execution, from proposal preparation to data analysis.
  - Support US users with telescope application
  - Support NOAO TAC with technical reviews as needed
  - Support TAC-approved programs
  - Liaise between Gemini and the US PIs in preparation for the ITAC to facilitate scheduling
  - Support post-data acquisition analysis
  - Inform the US community of Gemini opportunities

NGO activities

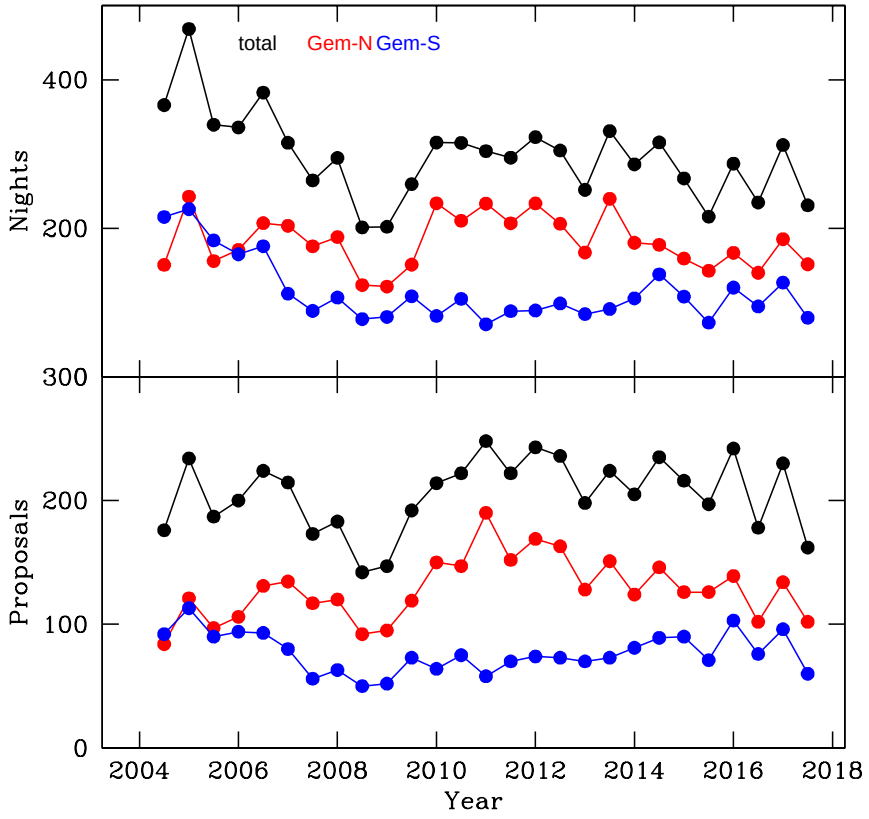
NGO supports post-data acquisition activities

NGO does not provide Phase II support, which is done by Gemini

Recent and current activities

- Update the GMOS Data Reduction Manual
- Update the NGO webpages with community-wide news, tools, and information useful for US Gemini users (and all users)
- A draft MIRI data reduction manual is now available
- AAS Mini-workshop series continues. January 2017 *Mining Data Archives* has been successful (see extra slides, including Gemini plans)
- 2018 AAS Mini-workshop is being planned at this time
- TAC support
- Supporting Phoenix as Gemini-S visitor instrument in 2017A/B

Gemini submission to NOAO TAC



These numbers exclude NOAO Surveys, and LLP Gemini programs.

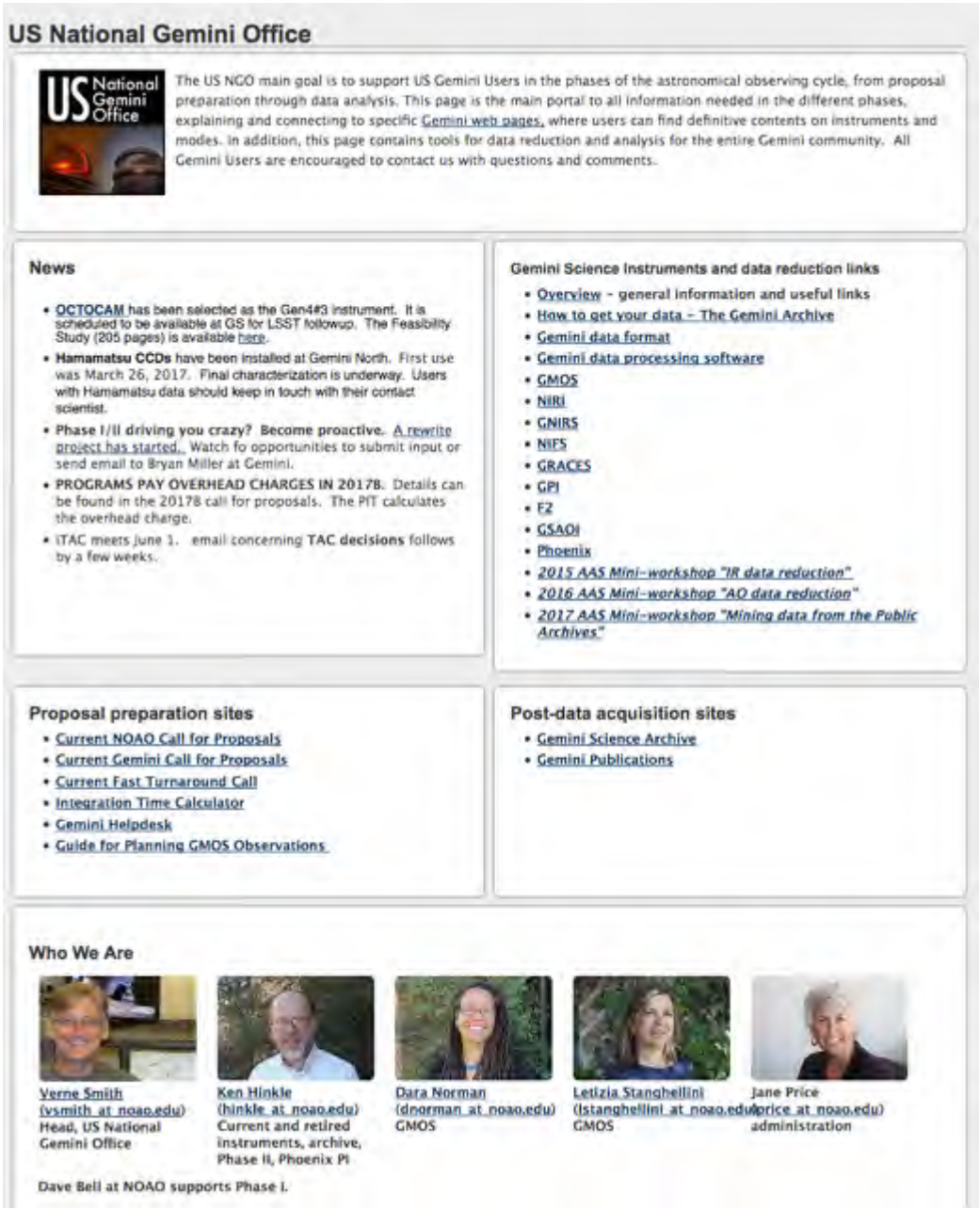
Submission has remained healthy at both telescopes.

NGO webpage

The NGO webpage is a gateway for all Gemini users

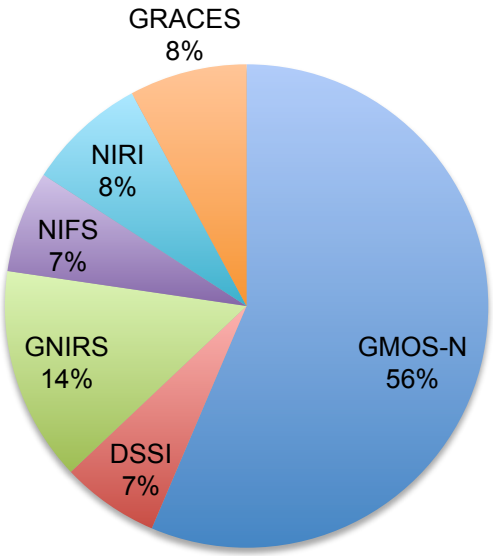
It is updated bi-weekly

It includes links to all data analysis tools available to date



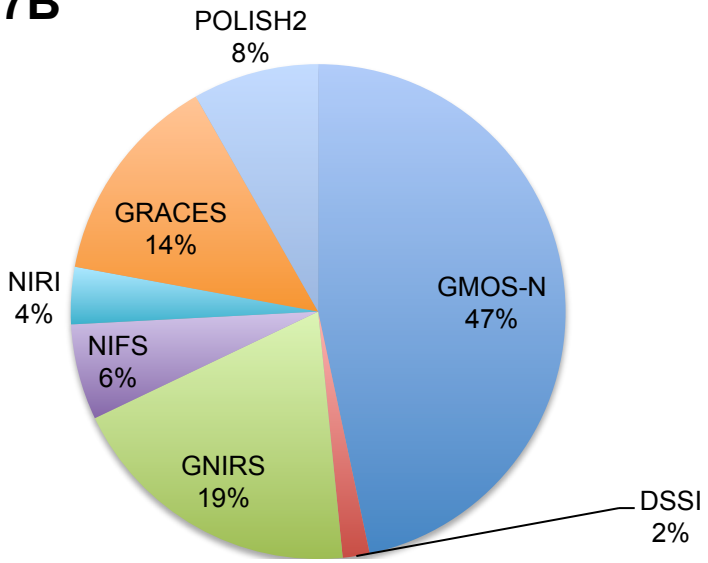
Gemini-N
Instrument requests (by nights)

2017A



Regular NOAO TAC
New programs only
(excluding Subaru exchange,
LLP, surveys, FT, DD)

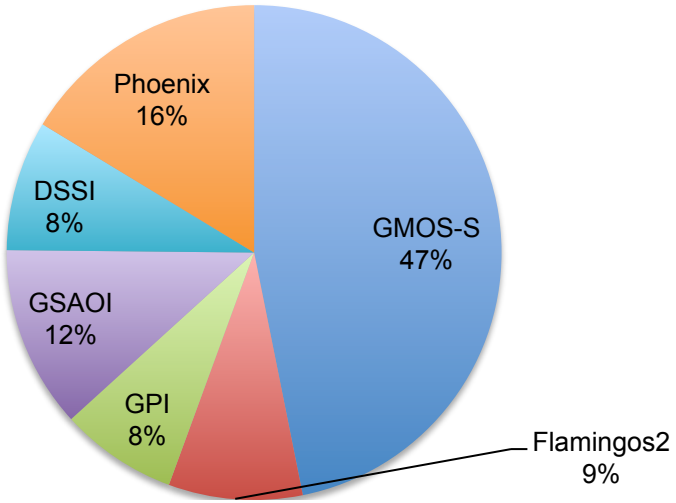
2017B



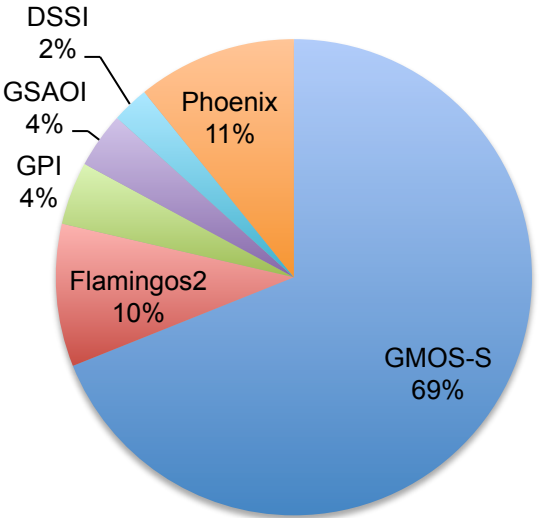
Gemini-N
Instrument requests (by nights)

Regular NOAO TAC
New programs only (excluding LLP, surveys, FT, DD)

2017A

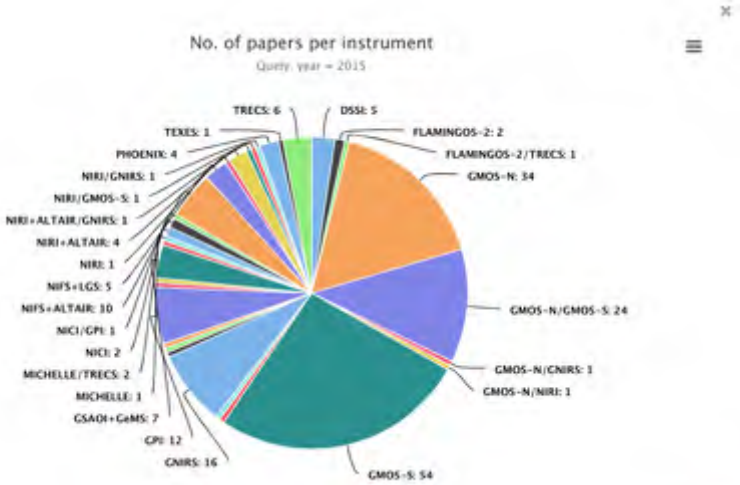


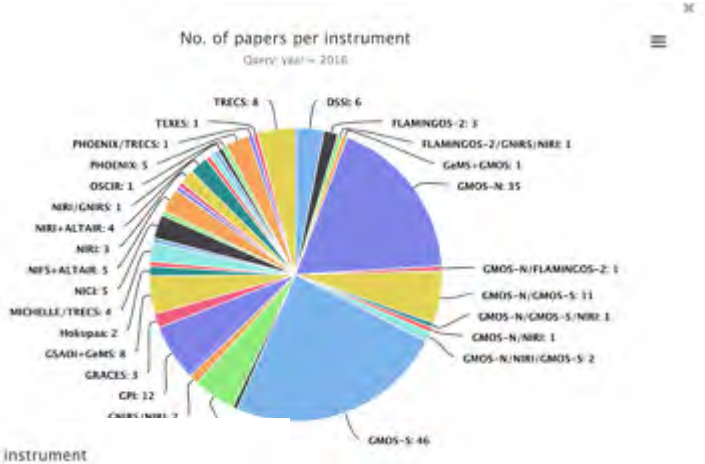
2017B

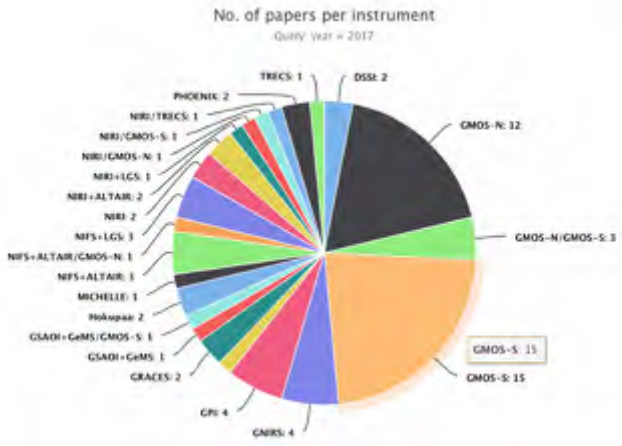


Publications from NOAO TAC allocated programs

- The number of US Gemini refereed publications has remained constant though the last few years (2015: 103; 2016: 90; 2017: 29 so far)
- Graph shows paper by instrument in 2015, 2016, 2017







Backup slides

2017 AAS Mini-workshop on *Mining Data Archives*

(1) Andy Adamson and Andre-Nicolas Chene: *Gemini Observatory Publications, Statistics, and Archive*
(2) Harry Teplitz: *Astrophysics Archives at IPAC*
(3) Scott Fleming: *Sailing the Archival Seas with MAST*
(4) Knut Olsen: *The NOAO Data Lab Project*

UC Meeting, Tucson, May 16-17, 2017

Gemini Observatory Publications, Statistics and Archive

A Adamson

A-N Chené

Abstract

We describe the statistics of Gemini refereed publications, including relative productivity and impact of instruments and observing modes, and overall statistics such as the total publication count. We identify factors which may influence the probability of a publication emerging from a given observing program. At present, only a small fraction of publications arise purely from archival data. We present some of our plans for post-observing community support, and solicit input on various options for increased productivity in archival research.

Notes in italics on each slide are the comments added to the information in the course of the talk.

Outline

Show some of the more interesting statistics relating to publications from Gemini observations - AJA

Show some of the ways in which we may be able to increase the publication rate, and solicit input on options - ANC

GEMINI OBSERVATORY

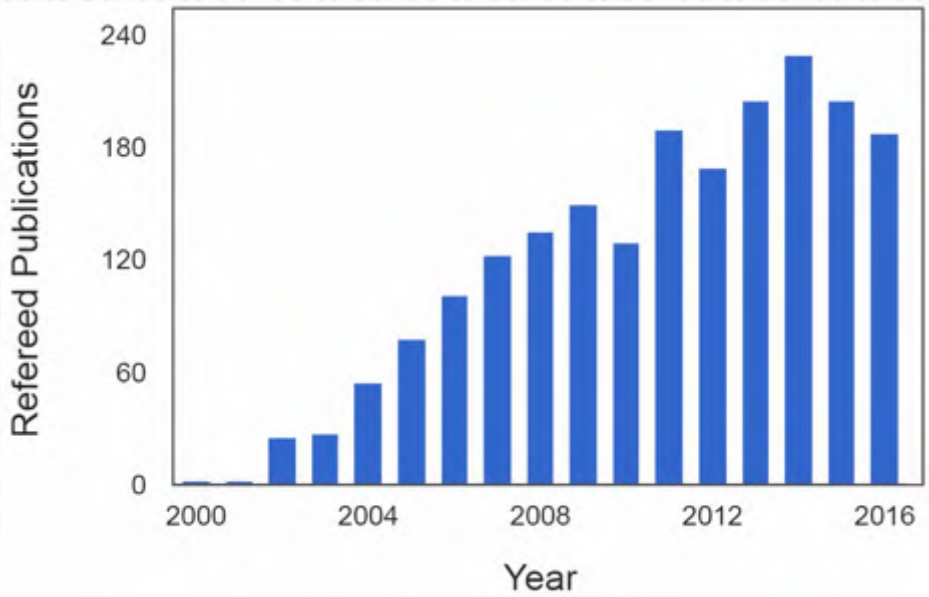
Exploring the Universe, Sharing its Wonders

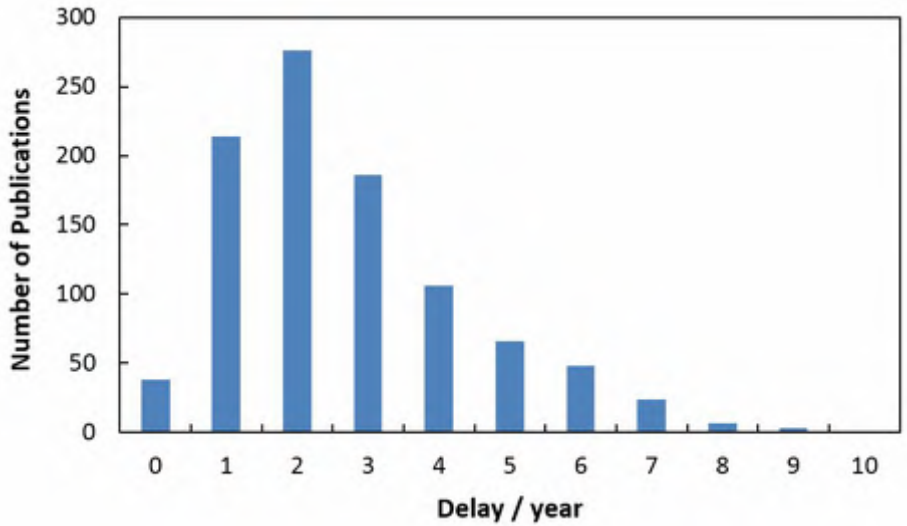
Publication Statistics

Overall Statistics

Publication count has now plateaued - about 200/yr, or a couple of papers per week at each telescope.

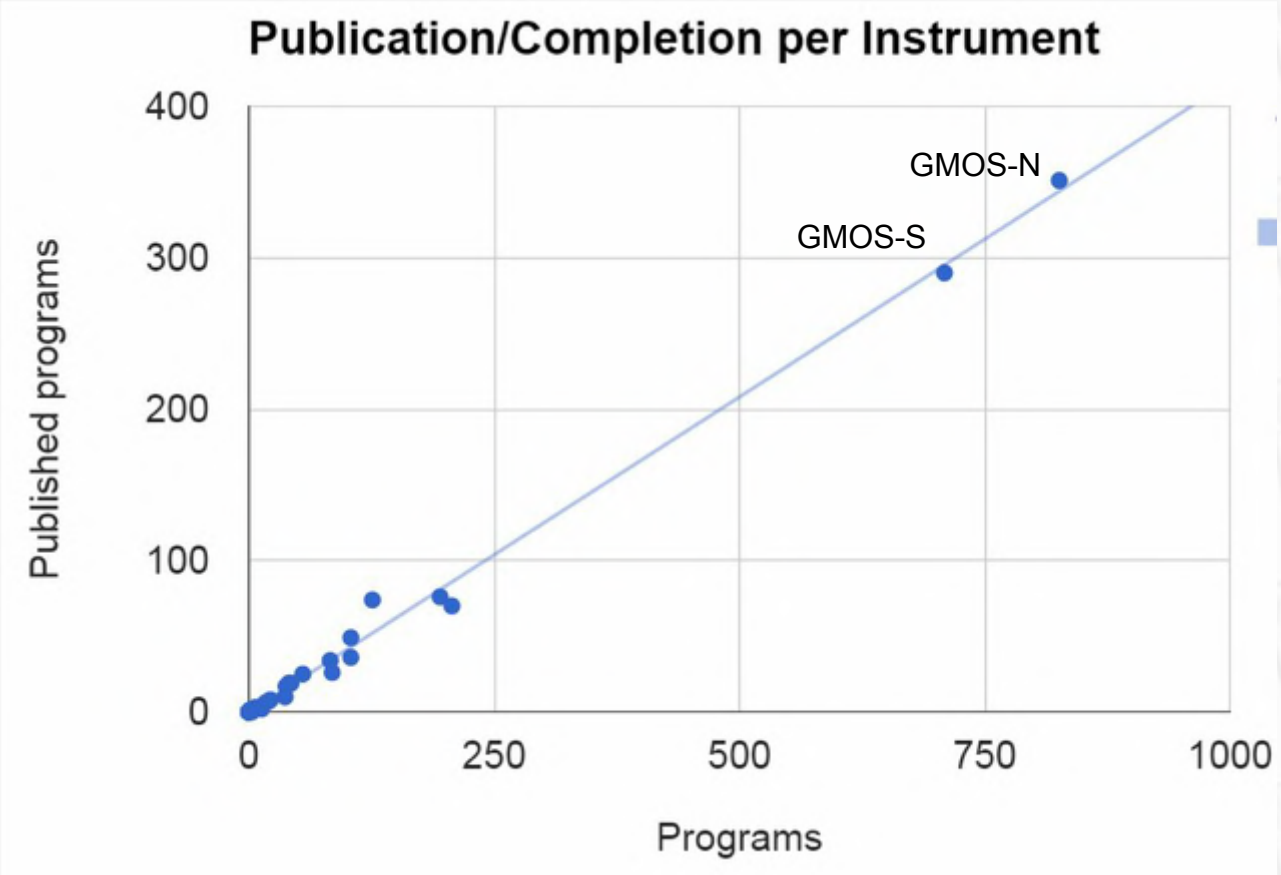
Lead time is about 2 years.





Noted in discussion: Gemini PIs across the whole partnership publish about two papers a week per telescope. Looks like we may now be at a plateau, and the influence of Large/Long programs is yet to be seen.

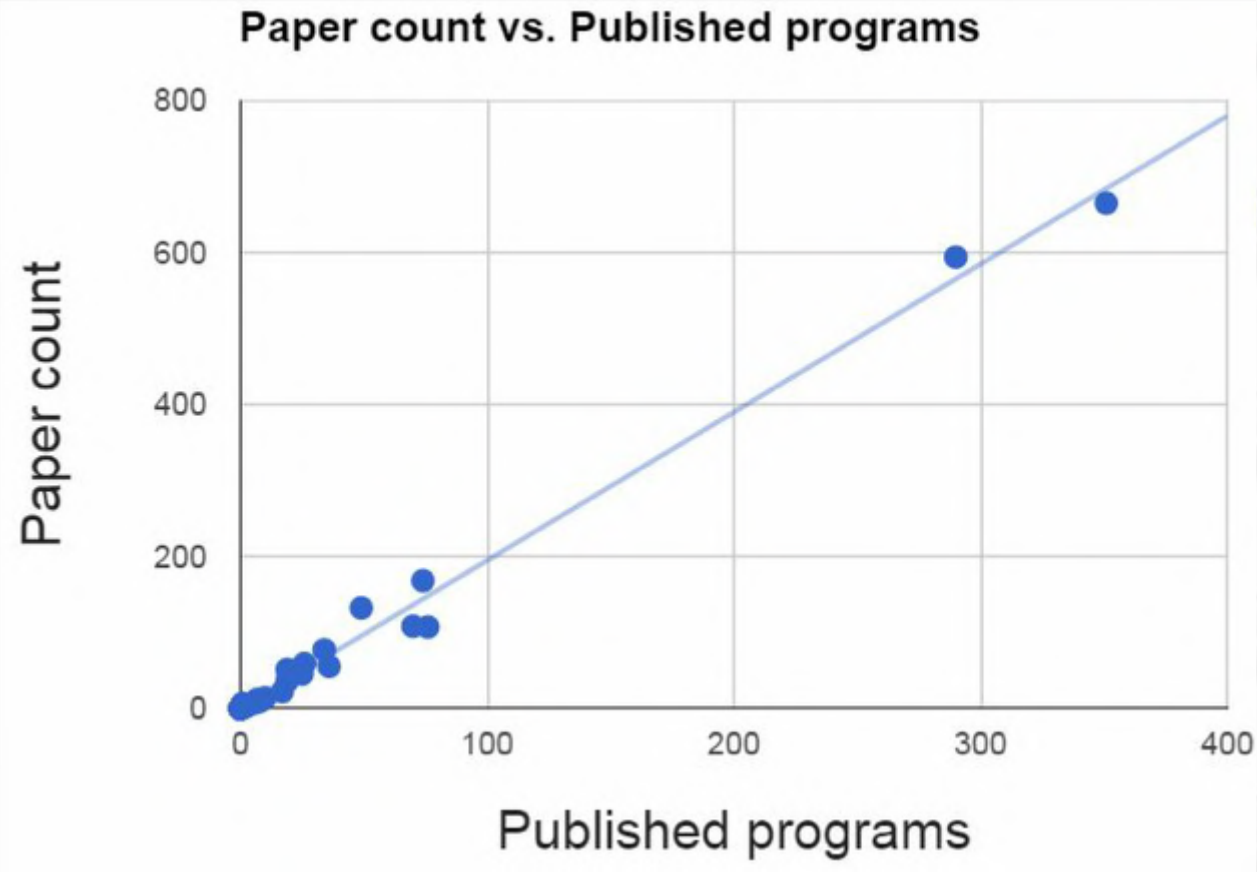
Publications vs Instrument



- >80% complete programs
- 2005-2014
- Each point is an instrument
- Single-instrument papers only

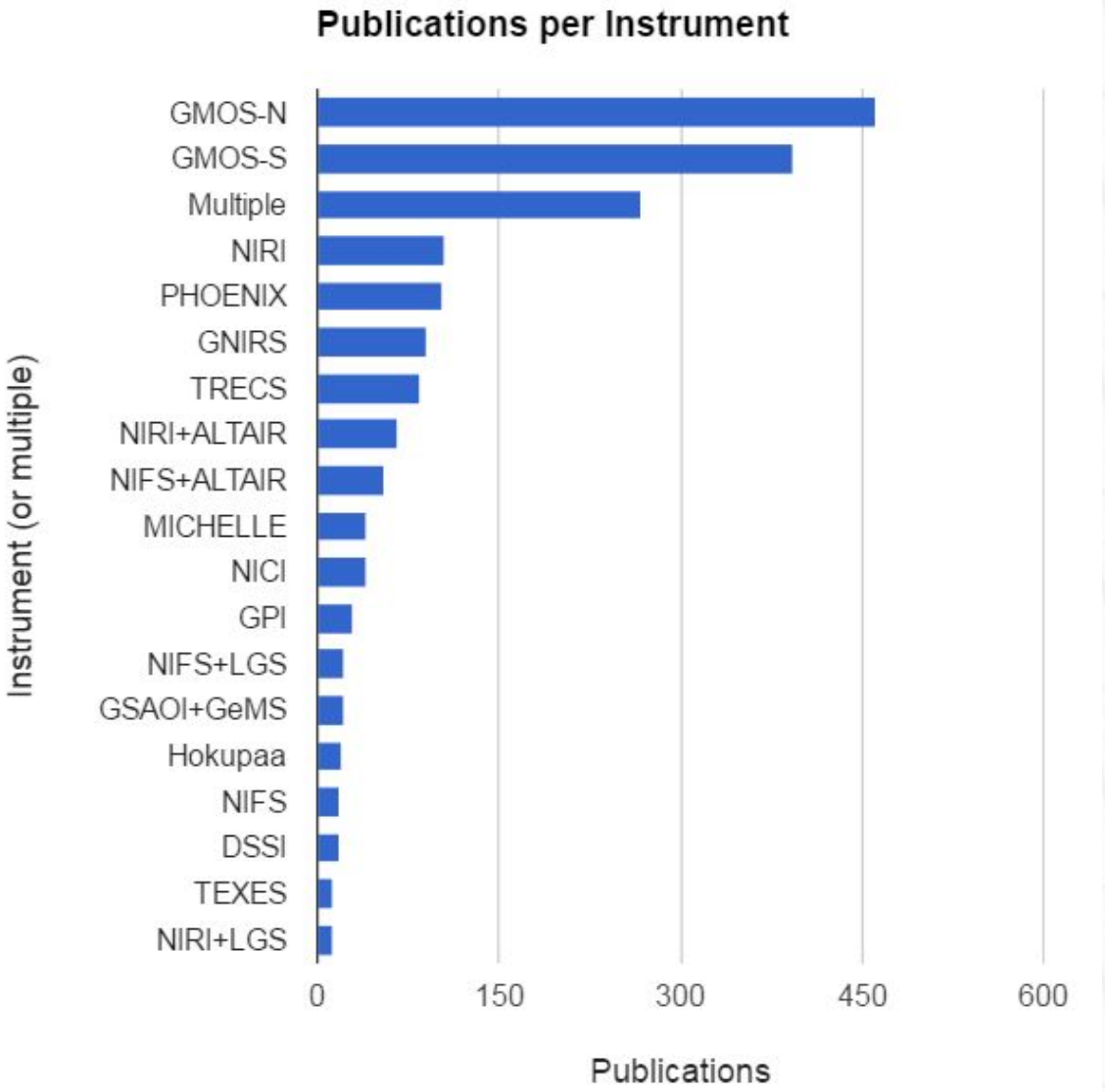
Noted in discussion: There's really great variation between instruments. All lie on the same line, with a slope of about 40-50% (meaning there's room to grow in terms of publication count).

Multiple publications



- A program that publishes, tends to publish twice
- Again, instrument-invariant

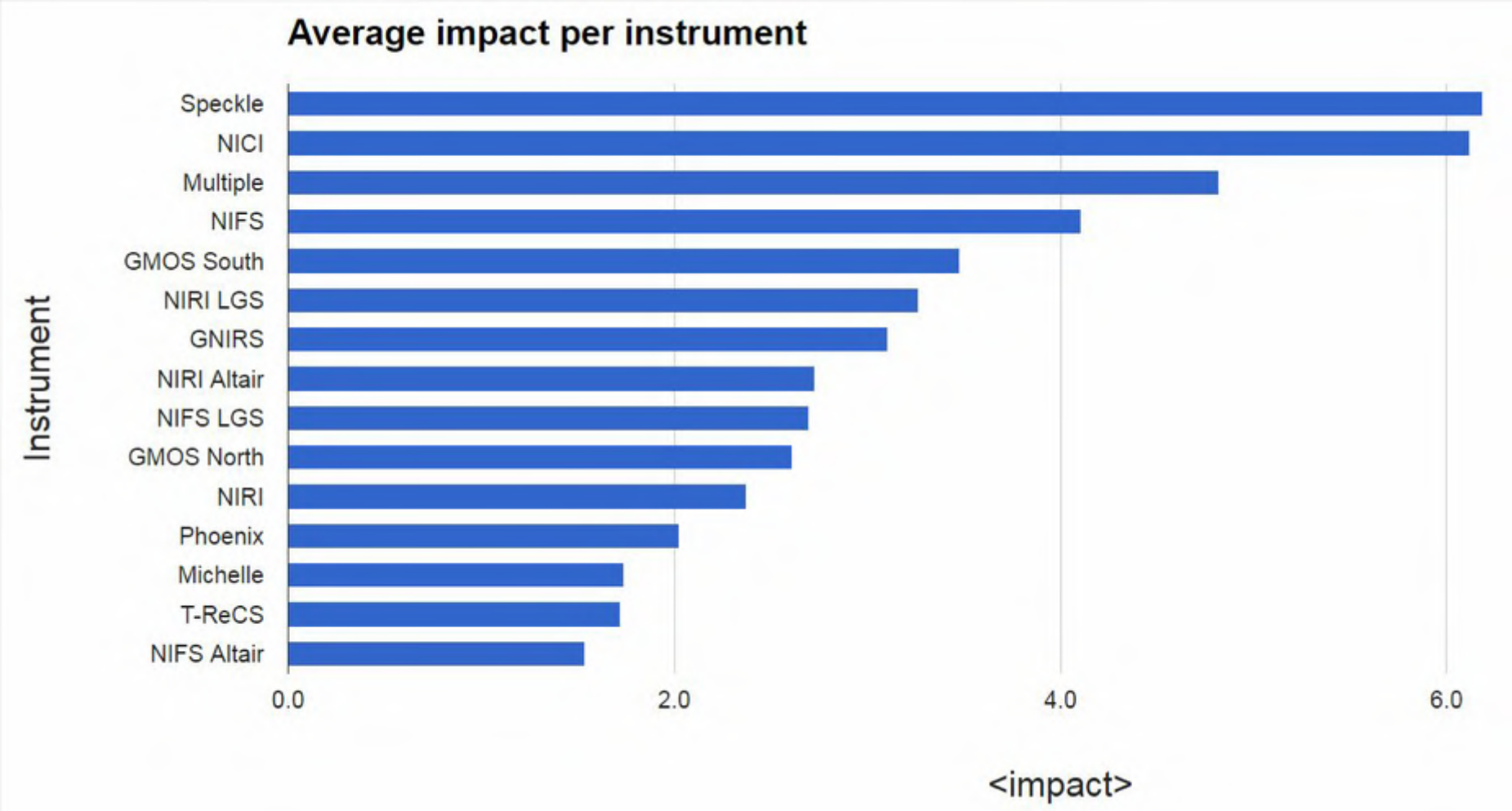
Publications by instrument



2005-2014

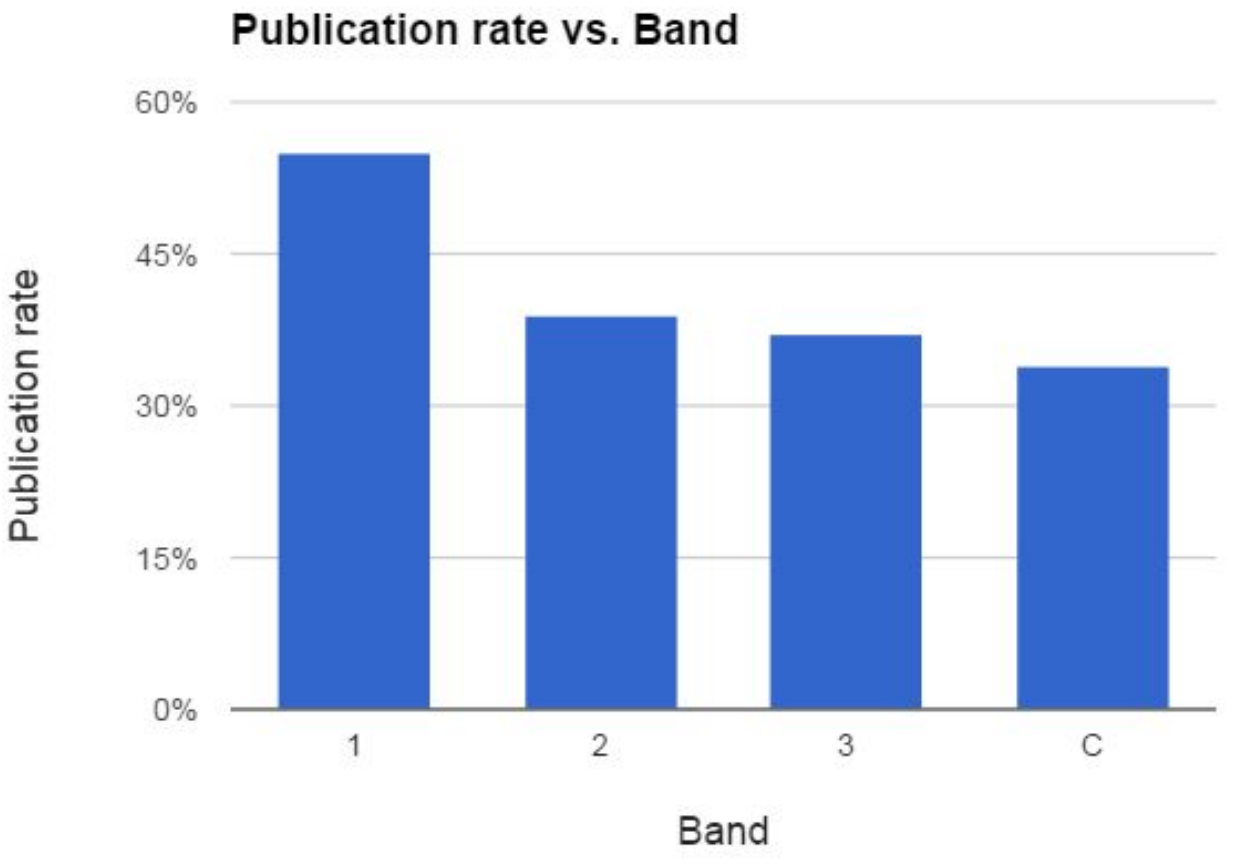
Note added in discussion: GMOSs being at the top is not surprising (they dominate the program count). "Multiple" instrument publications coming just below that is more interesting, probably reflects the many possible combinations of more than one instrument, and programs over more than one semester contributing to a publication.

<Impact> per instrument



Note in discussion: interesting that the cheapest instrument ever on the telescope has the highest current impact (but remember this is a snapshot). Also the fact that NICI - another exoplanet instrument - comes second shows that this chart is about subject fields, not instruments per se.

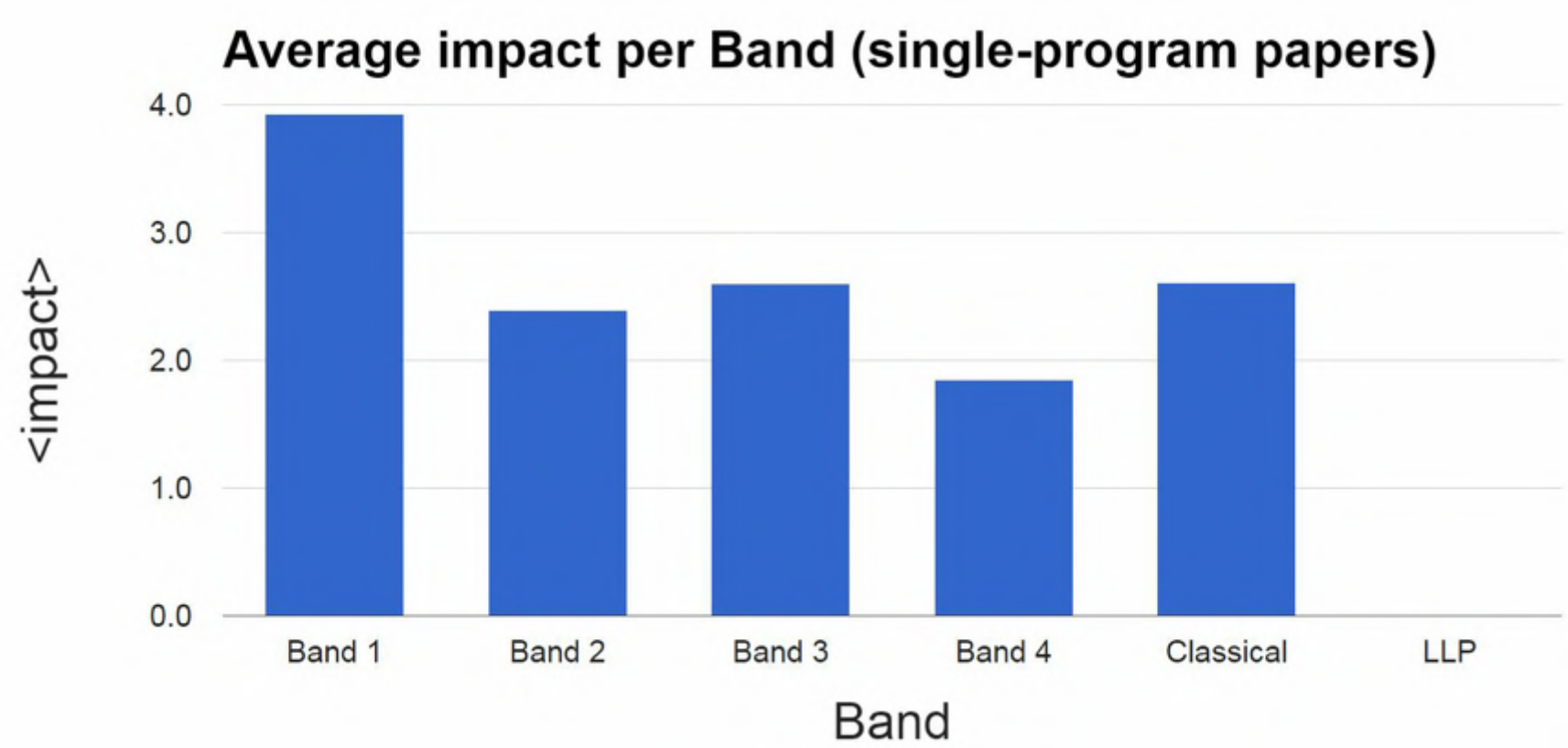
Publications vs band



- >80% complete programs
- 2005-2013

Note added in discussion: since this is all to do with programs at the same minimum completion rate, the indication is that the TACs can spot a winner.

<Impact> vs band



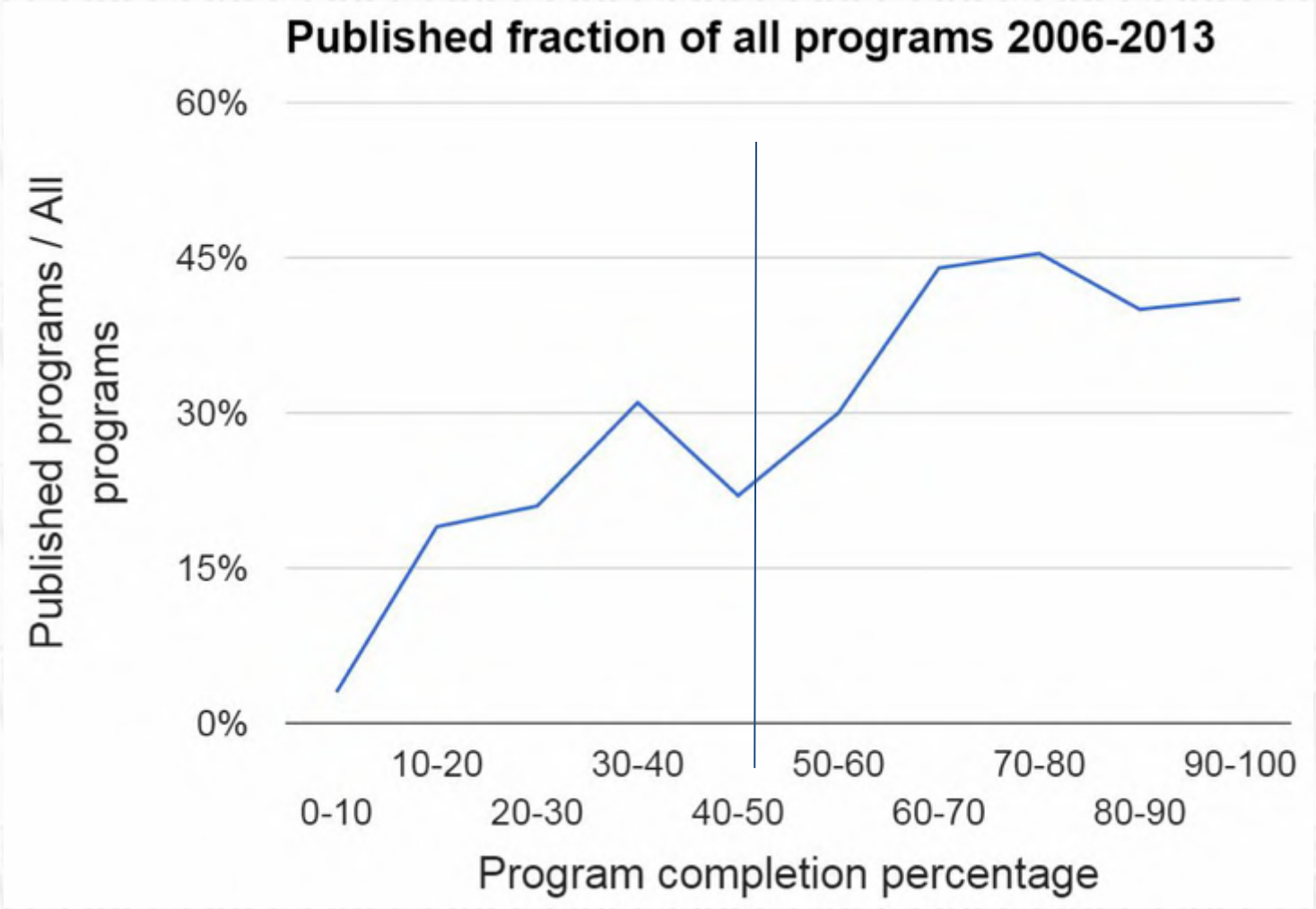
Single-program papers only, programs 2005-2015. Joints included.

Impact data from 2000-2014 publications. Updated Oct 2016

Nett coverage: publications between 2005 and 2013.

In discussion: within Band 1, Joint programs have by far the highest impact when published.

Publication vs completion



Note in discussion: once a program gets above 60-70% complete, the chance that its data will appear in the literature does not increase. This point has been discussed by Gemini and its Governance and has affected queue execution policy to some extent.

Feedback from Users

Data Gathering (NGOs):

2013-2015, with two-year lead time

Main Themes (consistent across partnership):

- Low priority data set <- longer programs
- Circumstances changed <- fast turnaround
- Postdoc left <- fast turnaround
- Didn't detect the target

Note in discussion: the most typical responses were not what had been expected. Changes in our offered observing modes may address some of these, as indicated in red.

Archive Publications

Papers emerging from purely archival data amount to only a few percent of all Gemini publications.

Gemini Science User Support, Data reduction and Archive possibilities

A-N Chené

Science User Support

- 4.6 FTE working on DR software
- 1 FTE focused on support

Avenues:

- Helpdesk
- Contact Scientists
- DR Forum
- Contacts page
  - Sus_inquiries
  - NGO contact info

Our Archive

So you have seen it at least once

https://archive.gemini.edu/

Reducing archive data

Two possible approaches:

- People can figure it out…

- Ain't nobody got time for that!!!

Resources currently available

Gemini reduction packages

- Gemini IRAF packages
- AstroConda!
  - set of astronomy-related packages for Anaconda
  - testing channel is now available
  - main Astroconda channel at STScI later this month

(One can contact James Turner from Gemini for any question about AstroConda.)

Reduction cookbook

http://ast.noao.edu/sites/default/files/GMOS_Cookbook/

But we could do more!!!

- Publishing more cookbooks
  - That is on the way (F2)
- Adding reduced data to archive
  - Ready for GRACES and GNIRS XDmode!
  - Reduced data provided by PIs
- Making “pipelines”
  - Would require a tremendous additional effort

GEMINI OBSERVATORY

Exploring the Universe, Sharing its Wonders

Thanks!

Questions?

Comments on priorities?

Caltech

Infrared Processing and Analysis Center

ipac

Astrophysics Archives at IPAC

Harry Teplitz

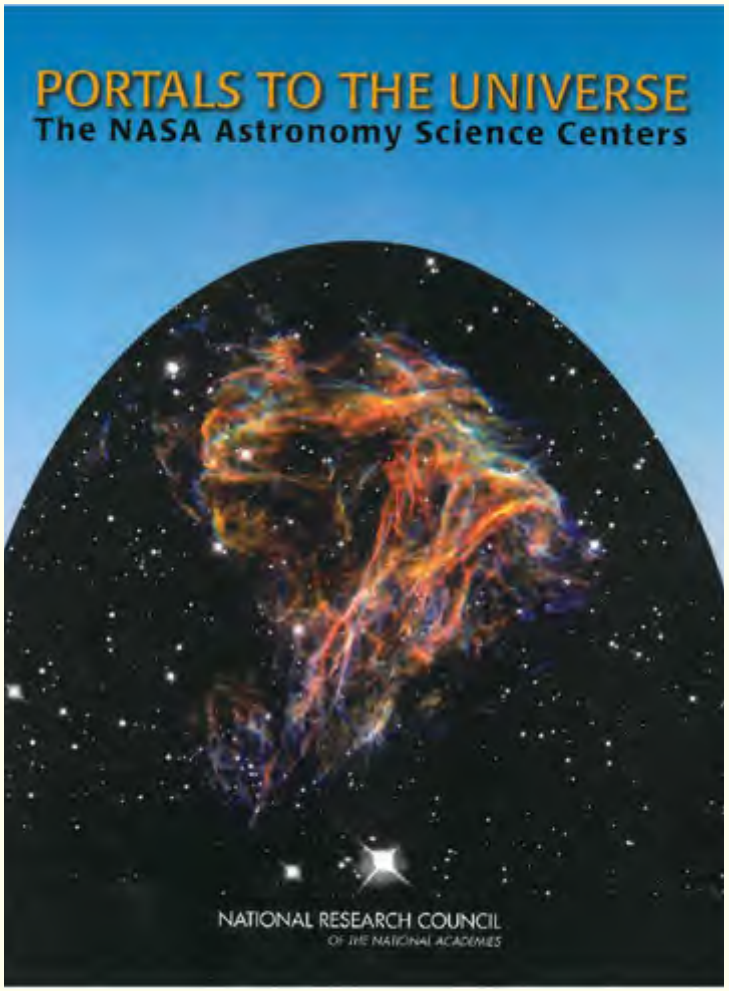
NASA's Commitment to Astrophysics Data Archives

- "NASA has regarded data handling and archiving as an integral part of space missions."
- "This support now provides the major return on the considerable investment the agency made… over the past 20 years."

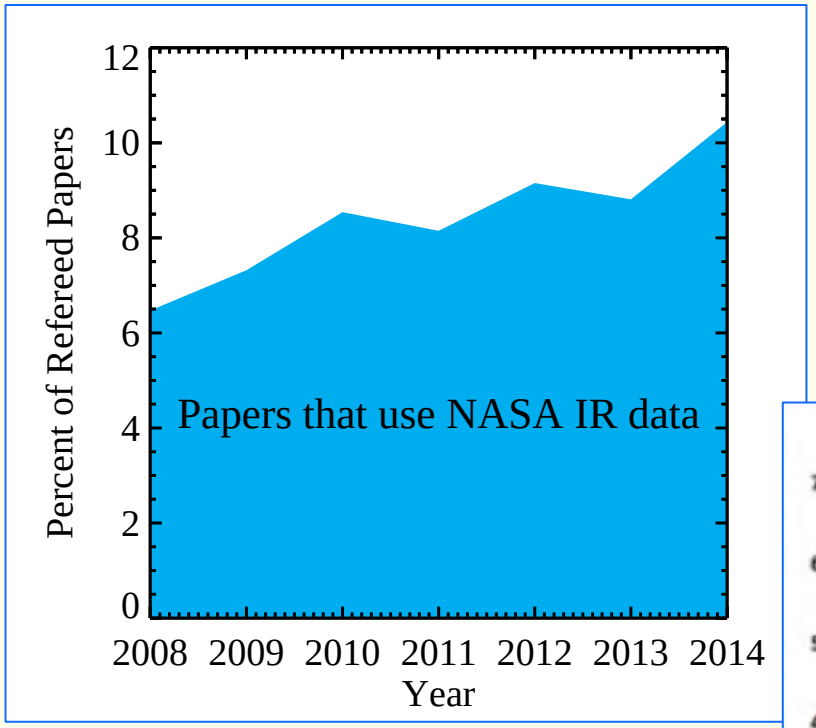
Infrared Processing and Analysis Center

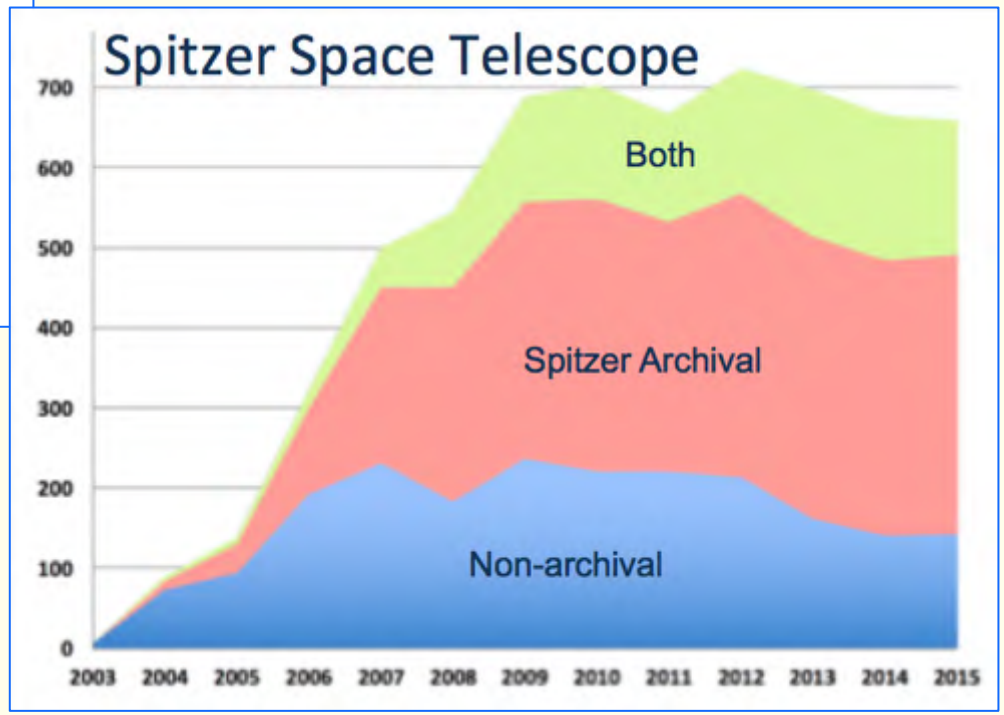
"A Sustainable Archive"

- Continually facilitates production of new scientific results
- Has a strategic goal to enable more and better science
- Contains high-quality, reliable data
- Provides simple and useful tools to a broad community
- Provides user support to the novice as well as to the power user
- Has many diverse uses (and users)
- Adapts and evolves in response to community input



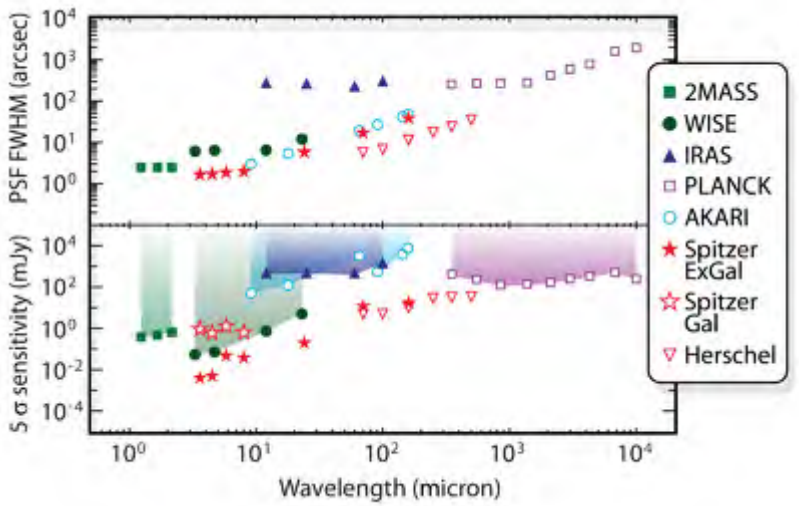
Archives Double the Number of Papers from the Observatory





IRSA : NASA's IR/sub-mm archive

- IRSA ensures the legacy of NASA's "golden age" of IR
  - *Enable research that has not yet been envisioned.*
  - *Priorities set by missions and the community*
  - *Support future flight missions*
- IRSA is continuing rapid expansion
  - *Since 2011, holdings more than doubled (now > 1 PB);*
  - *# table rows increased by factor of 15 (>100 billion)*
  - *Almost 40 million queries in 2016*



- *All-sky 20 photometric bands from 1 micron to 1 cm*
- *About 40% of approved ADAP programs involve analysis of IR data sets*

Spitzer WISE Planck Herschel 2MASS IRAS AKARI MSX SWAS IRTS BLAST SOFIA

NED : managing complex data sets

Overview – NED in a Nutshell

Published:
- Names
- (α.δ)
- Redshifts
- D_{Mpc}
- Fluxes
- Sizes
- Attributes
- References
- Notes

Contributed:
- Images
- Spectra

Derived:
- *Distances*
- *Metric sizes*
- *Luminosities*
- *Velocity corrections*
- *Cosmological corrections*
- *SEDs*
- *A_λ*

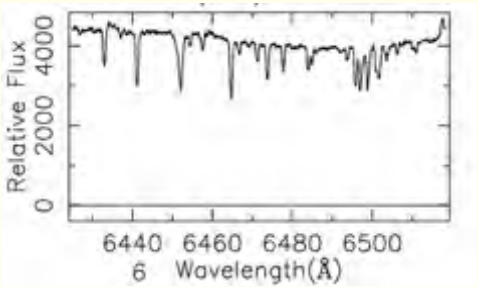
NED is where you find...

- Objects with z > 2.0 and available GALEX NUV flux
- Most precise z-independent distance measurement to M82
- SED, *spanning gamma-rays through radio*, for quasar 3C 279
- Spiral galaxies with stellar bars and Type 2 AGNs

Extragalactic papers have grown to 3,500 per year, with unique measurements for millions of objects

The Keck Observatory Archive

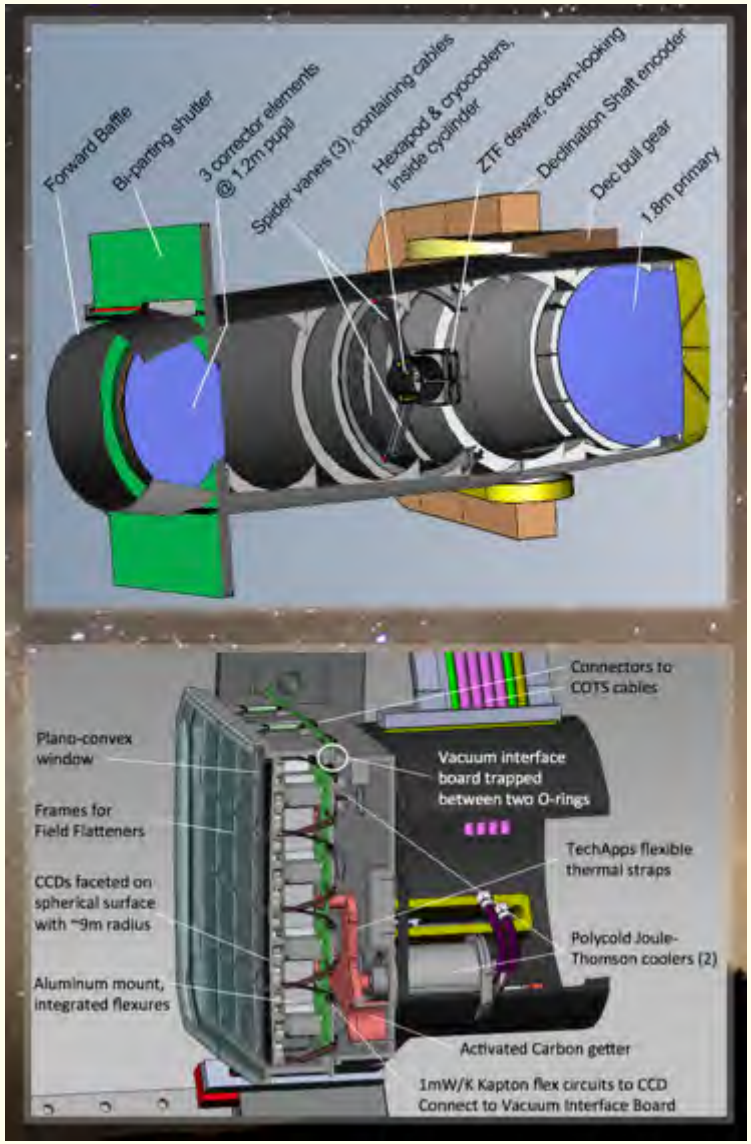
- NASA-funded collaboration between WMKO and IPAC/NExScI.
- Started with HIRES
- Systematic/automated capture of metadata ensures efficiency
- Now data from all ten instruments since their dates of commissioning
  - *decommissioned instruments Summer 2015.*
  - *proprietary period of at least 18 mo.*
- KOA creates browse products for three instruments by automating pipelines.



KOA creates extracted HIRES browse spectra for every order of each object raw frame. Shown: T Tau. (PI: Reipurth).

Infrared Processing and Analysis Center

Other IPAC Archives following the NASA model (non-NASA funded)



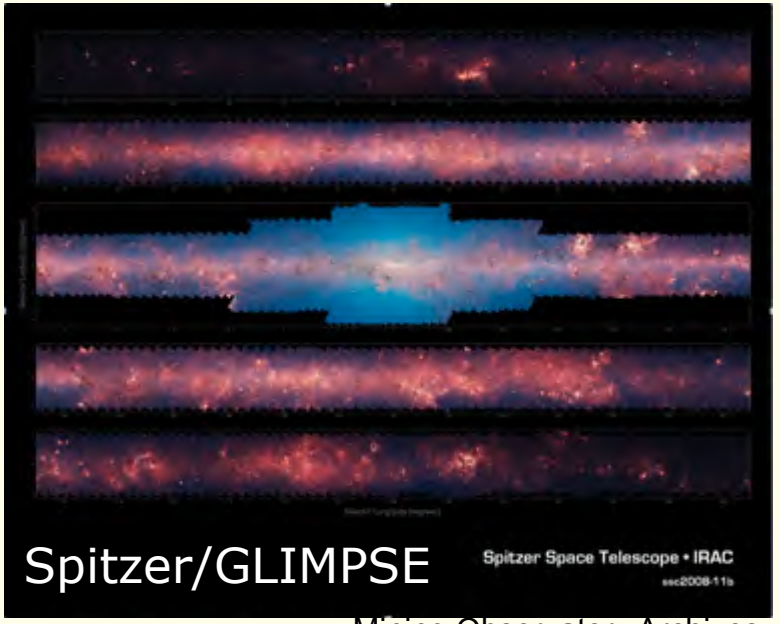
- **Palomar Observatory**
  - *Zwicky Transient Facility (2017+)*
  - *intermediate Palomar Transient Factory (iPTF; 2013-2016)*
  - *Palomar Transient Factory (2009-2012)*
- Fully automated wide-field survey with 1.2 m Oschin telescope
- Publicly accessible survey data products available at IPAC
  - *single frame exposures for selected regions of the sky,*
  - *source catalog files for those same regions.*

An Archive's Job

- Ingest new data
- Maintain/serve vital repository of irreplaceable data
  - *Support for observation planning*
  - *Resource for original science*
  - *High level science products*
- Enable cutting-edge research
  - *API and Virtual Observatory*
  - *User support by experts*
  - *New/enhanced services*

High Level Science Products

- Greatly enhance the science return of the archives
  - *Hubble Legacy HLSP are used 10x as much as typical pipeline products*
- Make complex data sets accessible to a wider audience of researchers
- Expand the use of large, coherent projects
  - *Herschel Key Projects*
  - *Spitzer Legacy and Exploration Science*
- Generated by the community or by the archive



Technical Synergy and Innovation

- IRSA implemented innovative indexing techniques for NEO/ WISE, optimized to meet the required use cases for database queries
  - *single position spatial searches (using a recursively subdivided triangular mesh)*
  - *simultaneous matching of large user-supplied lists of positions (using a file-based index outside of the database).*
- Challenges presented by WISE were used as opportunities to extend IPAC's capabilities
  - *ZTF will require databases that are at least an order of magnitude larger*

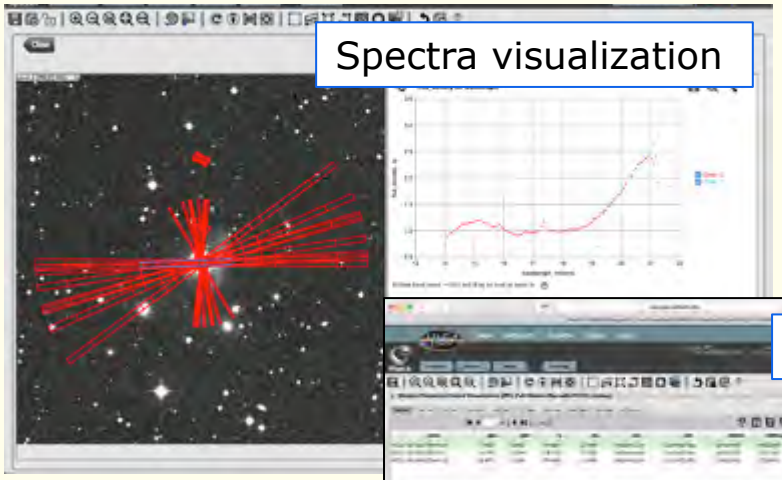
Science User Support

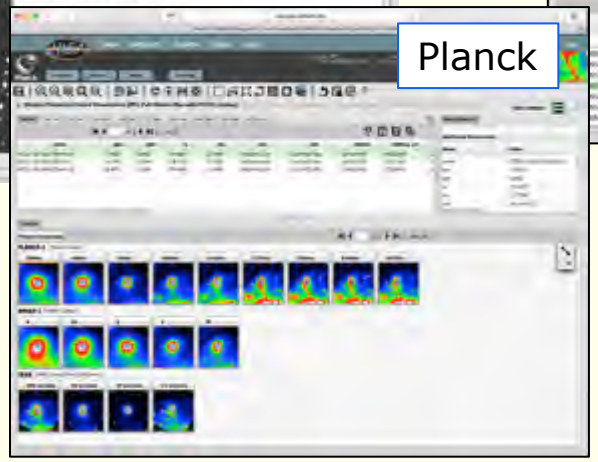
- *Helpdesk* some tickets are simple, others extremely complex
- *Documentation*
  - *tools/data releases*
  - *updates in response to tickets*
  - *Handouts*
- *Demos*
  - *Live (AAS, ADASS, DPS)*
  - *Video tutorials (IRSA has > 60 videos)*

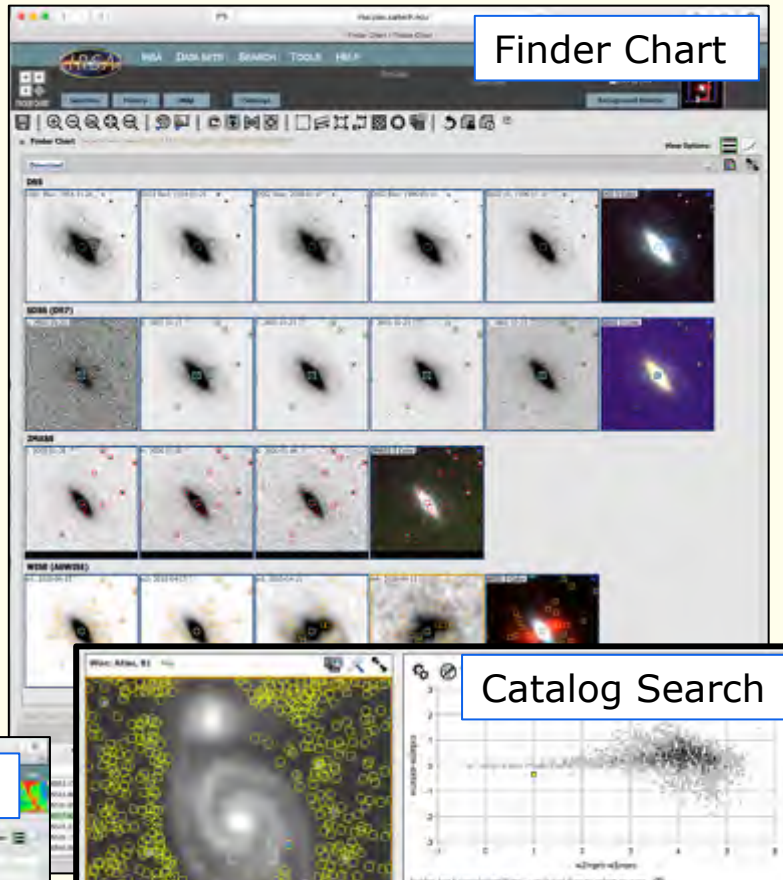
- The complexity of Science User needs increases with time.

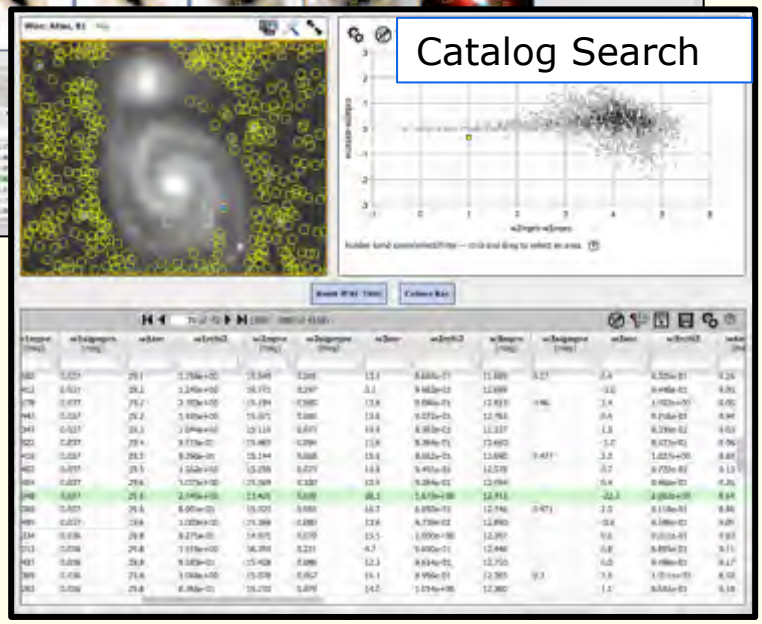
Data Exploration and Visualization Services

- Search & display can be tailored to various instrument/science contexts, using reusable visualization components
- Combine images, plots, tables, spectra
- Supports observation planning
- Firefly by IPAC







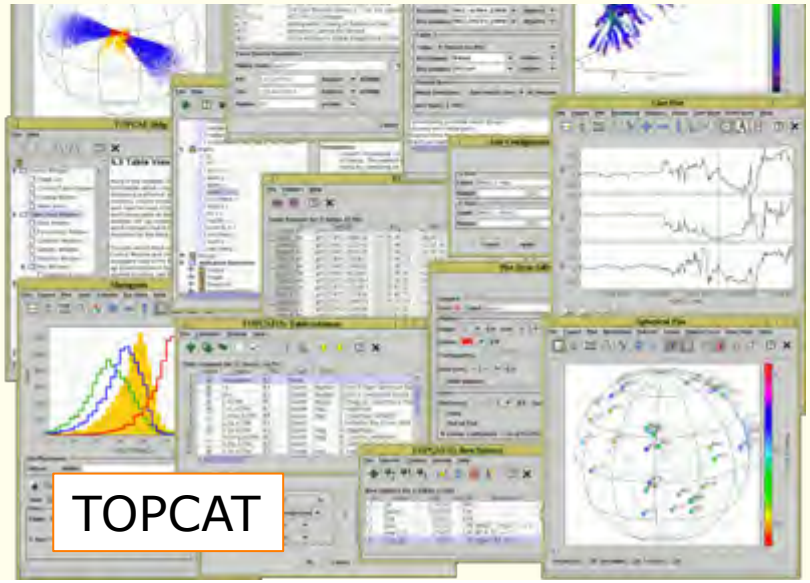


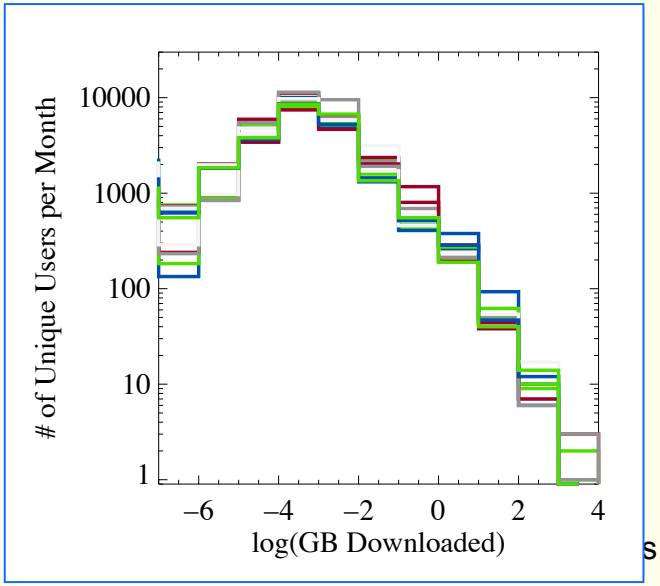
Infrared Processing and Analysis Center

ipac

VO broadens audience; API supports diverse users

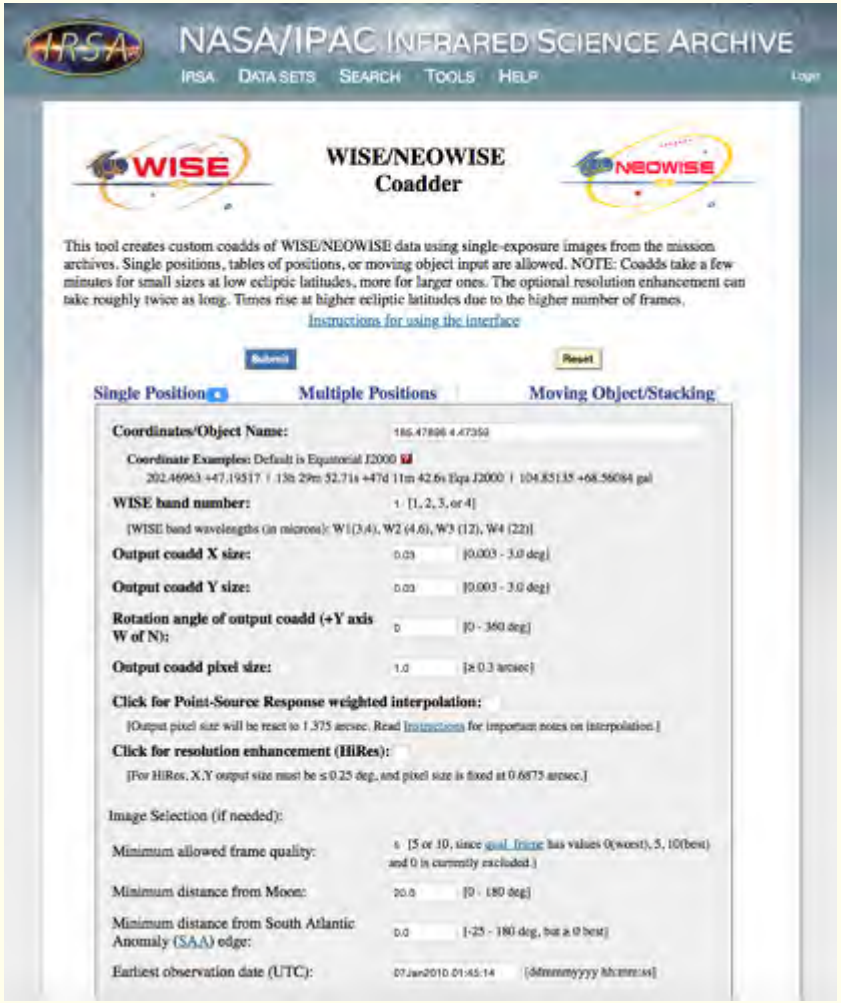
- Virtual Observatory
  - *Standardized protocols for interoperability between archives*
  - *Data discovery*
  - *Independent Tools*
- Application Program Interface
  - *Allows scripted access to archive data*
  - *Enables complex projects*





Data Analysis Tools

- Analysis "near the data"
  - *As data sets grow beyond local resources, researchers look to data centers to provide computing power and tools*
- Interactive tools
  - *WISE Coadder*
  - *Planck map making*
  - *IRAS tools*
  - *Spitzer imaging and spectra*
- Looking Forward
  - *"Big data means you can't move it all", suggesting the analysis must move to the data*

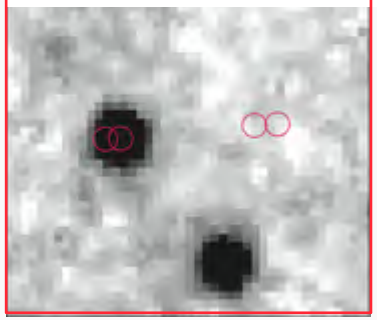


IR Science Highlights

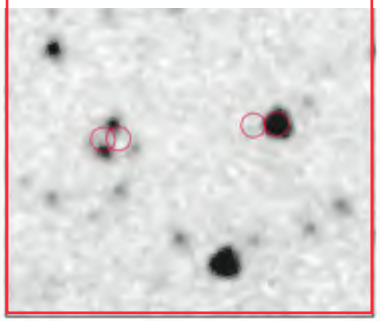
Infrared Processing and Analysis Center

Time domain

WISE 4.5μm (2010)

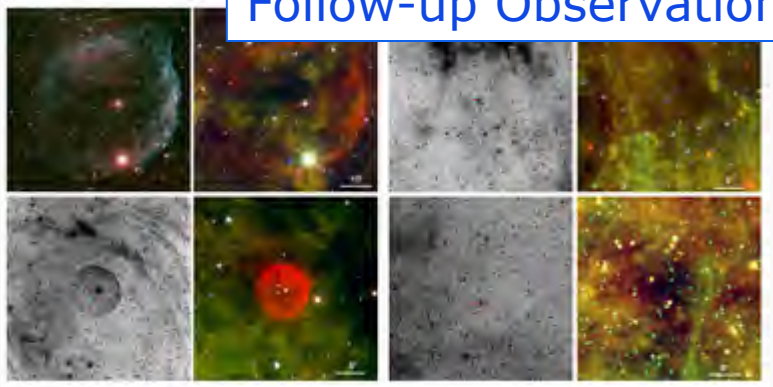


Spitzer 4.5μm (2014)

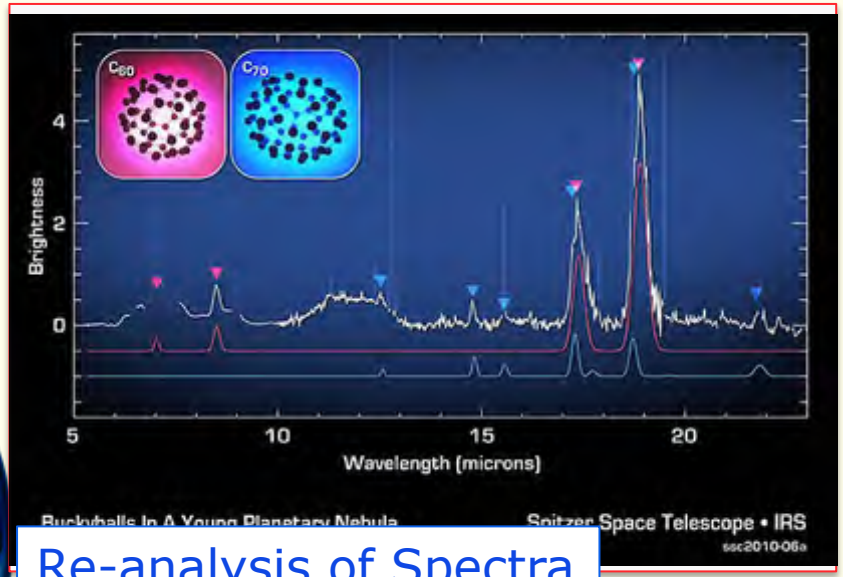


WISE+Spitzer discover the coldest brown dwarf (Luhman 2014, ApJL 786, L18)

Follow-up Observations

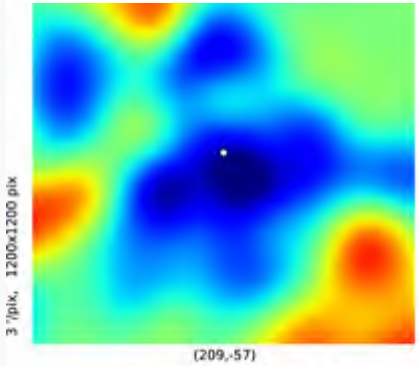


WISE morphological study of Wolf-Rayet nebulae, Toala et al. (A&A 2015, arXiv:1503.06878)

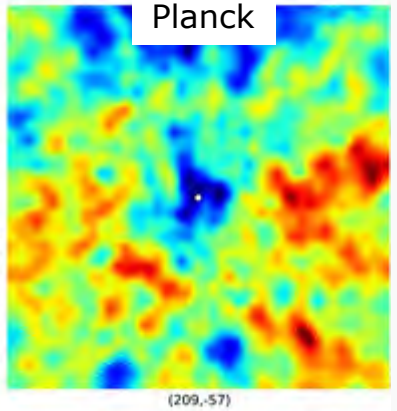


Re-analysis of Spectra

WISE+2MASS galaxies



Planck



WISE+2MASS+PanSTARRS data may reveal super-void in CMB cold spot seen by Planck; (Szapudi et al. 2015, MNRAS, 450, 288)

Combination of Surveys

Lessons Learned

- Long-term, stable archives greatly increase the return on observatory investment

- Robust support for both expert and novice users pays off
- User support by instrument experts is crucial
- Standardization of tools within an archive increases efficiency
- Integrity of science data as obtained must be maintained
- Interoperability between archives benefits everyone
- High level data products can expand the reach of large data sets

Spitzer image of infrared dark clouds

Sailing the Archival Seas with MAST

Scott W. Fleming

STScI

on behalf of the MAST team

Outline – Mining Archival Data

- Overview of MAST missions
- The MAST Discovery Portal
  - Cone Search
  - Advanced Search
  - Cross-Matching
  - Access to the Virtual Observatory
- MAST-led Enhanced Data Products
  - gPhoton: GALEX photon events
  - HSC: The Hubble Source Catalog

MAST Missions: Past, Present and Future

Legacy Missions


Astro


Copernicus


EPOCh


EUVE


FUSE


GALEX

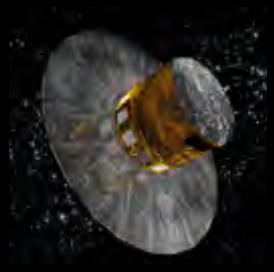

IUE


Kepler


ORFEUS


VLA - FIRST

Active Missions


Gaia


Hubble


K2

Pan-STARRS

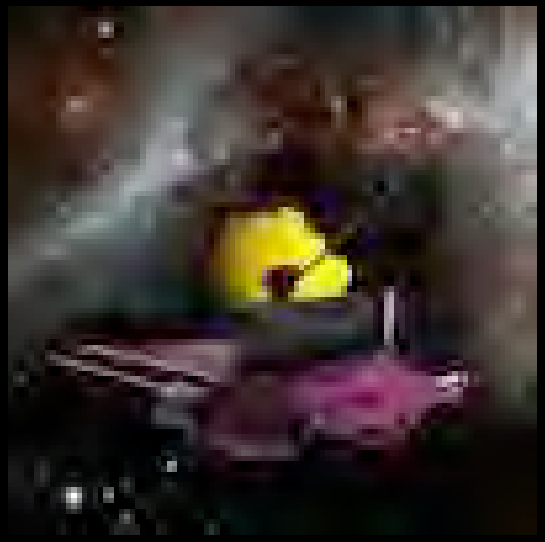

Swift UVOT


XMM OM

MAST Missions: Past, Present and Future

Upcoming Missions

TESS – Mar. 2018

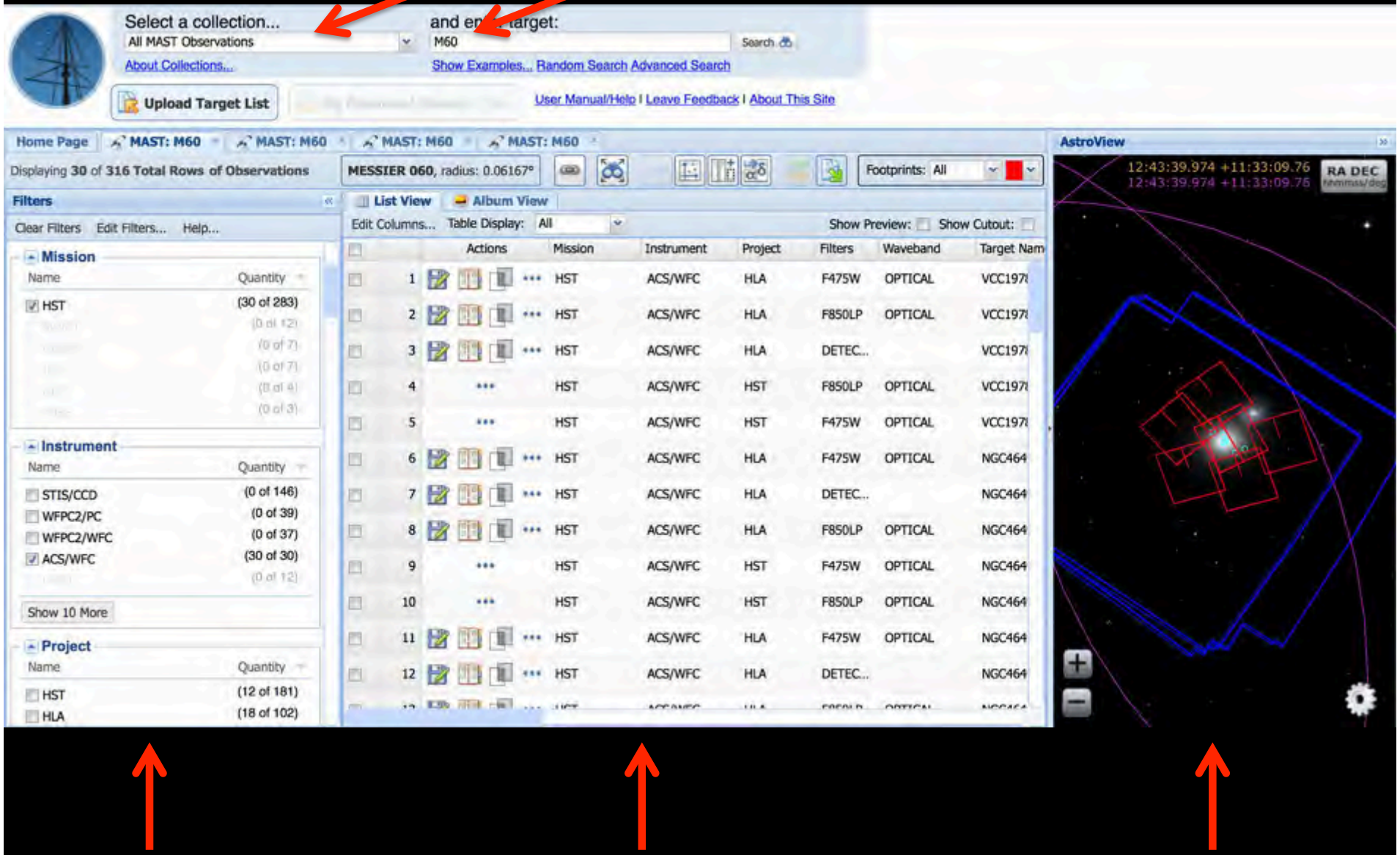


James Webb – Oct. 2018

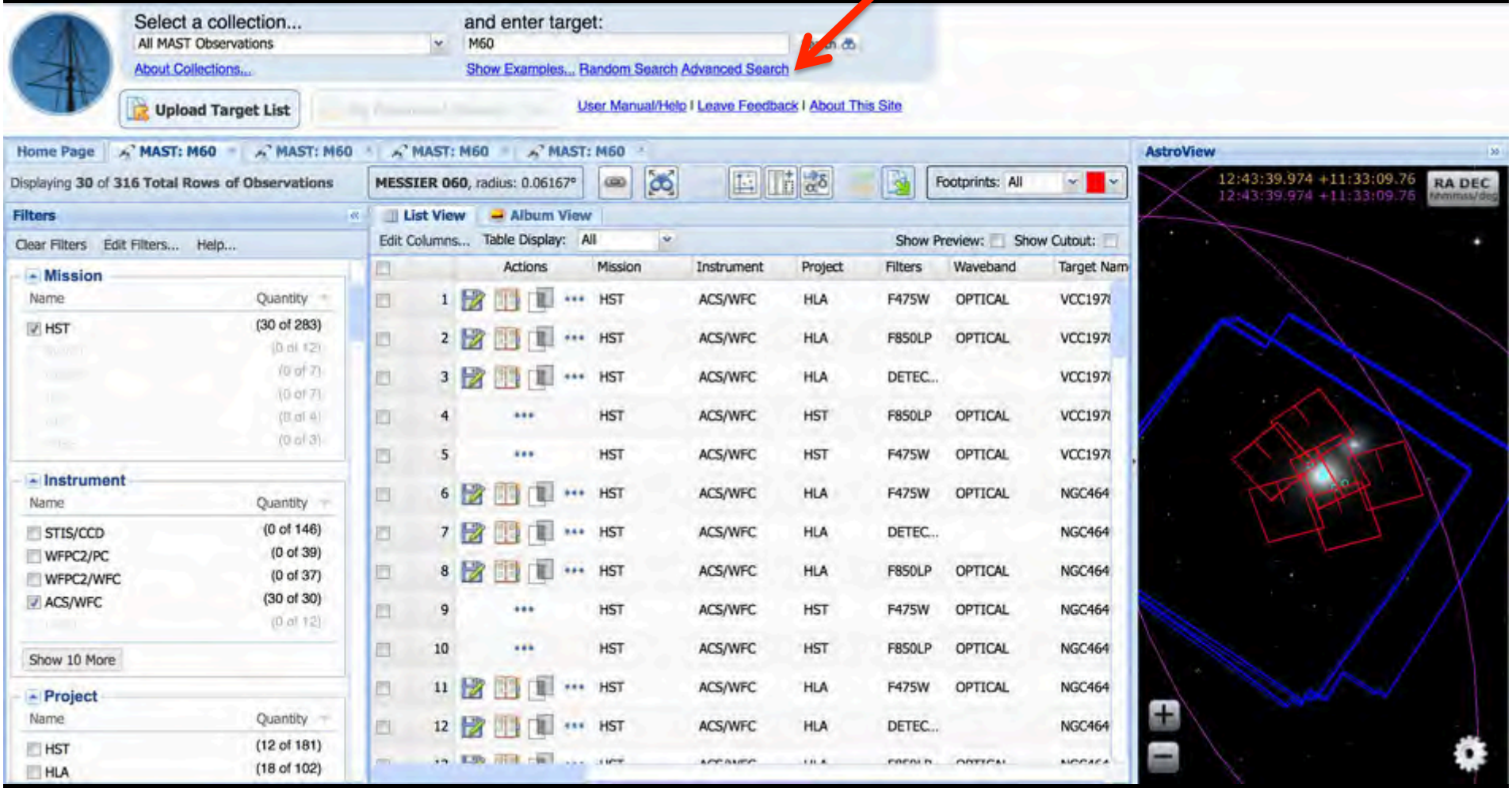
WFIRST – mid-2020s

Cross-Mission Discovery

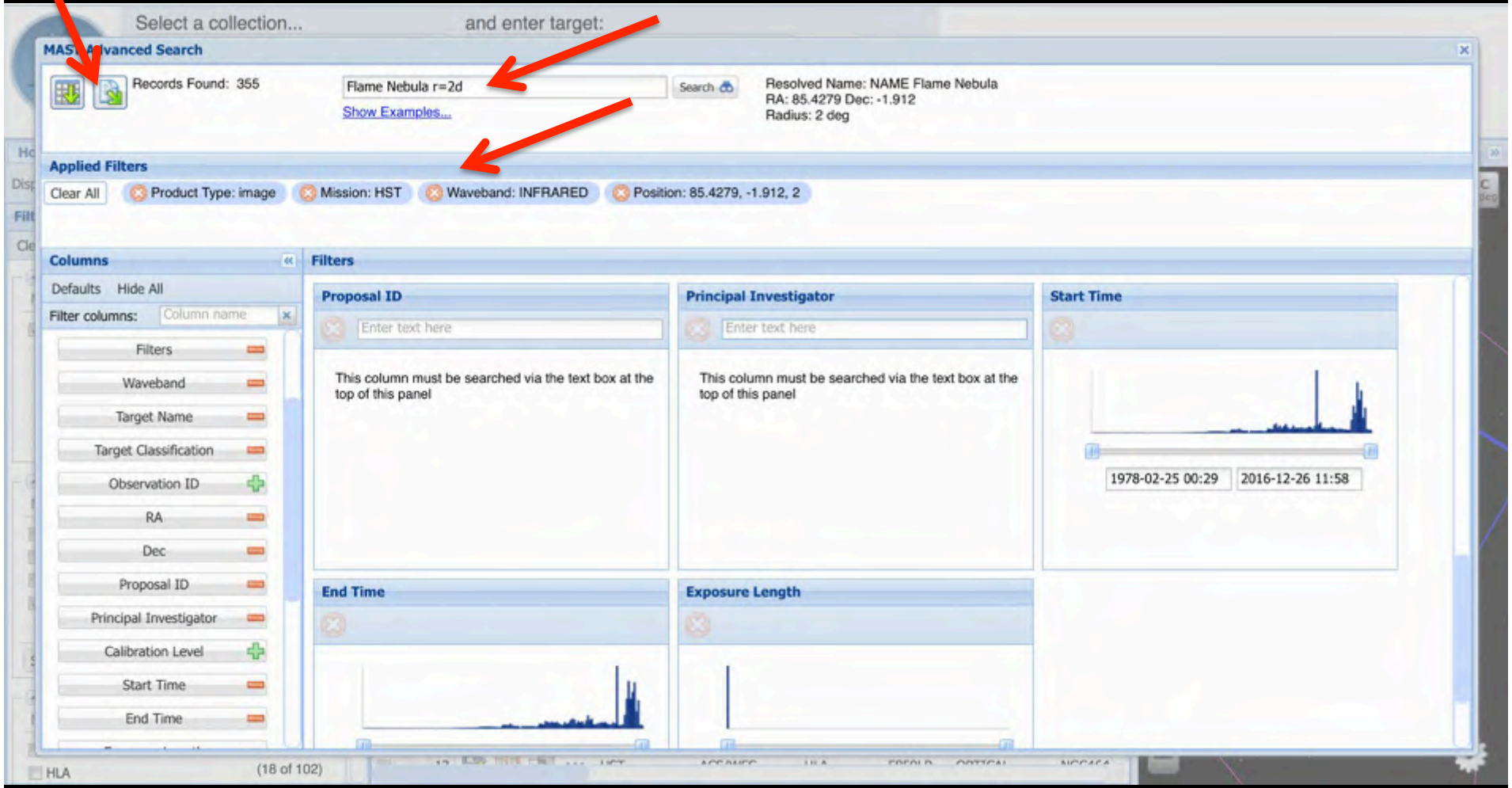
MAST Portal: Cone Search



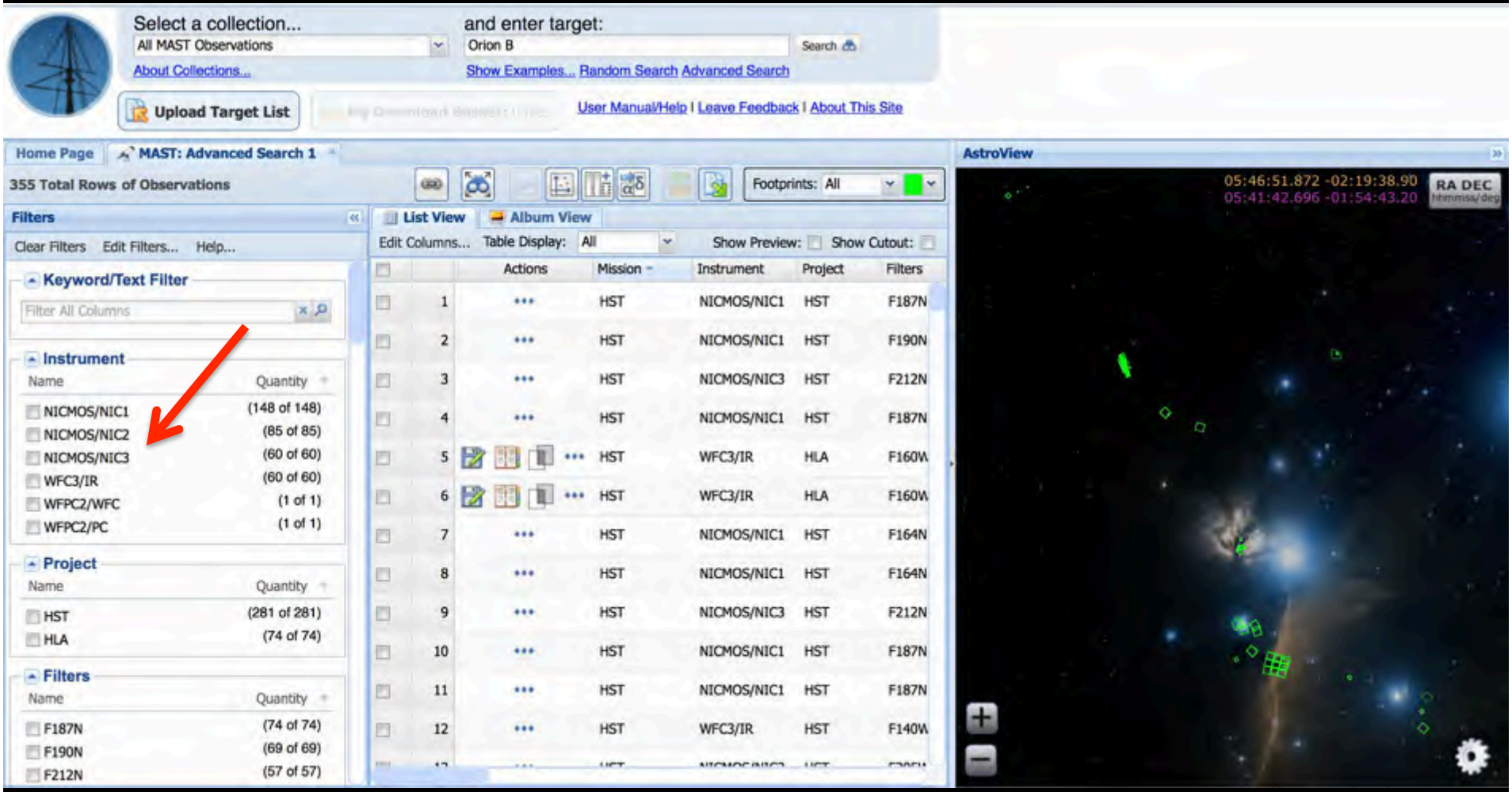
MAST Portal: Advanced Search



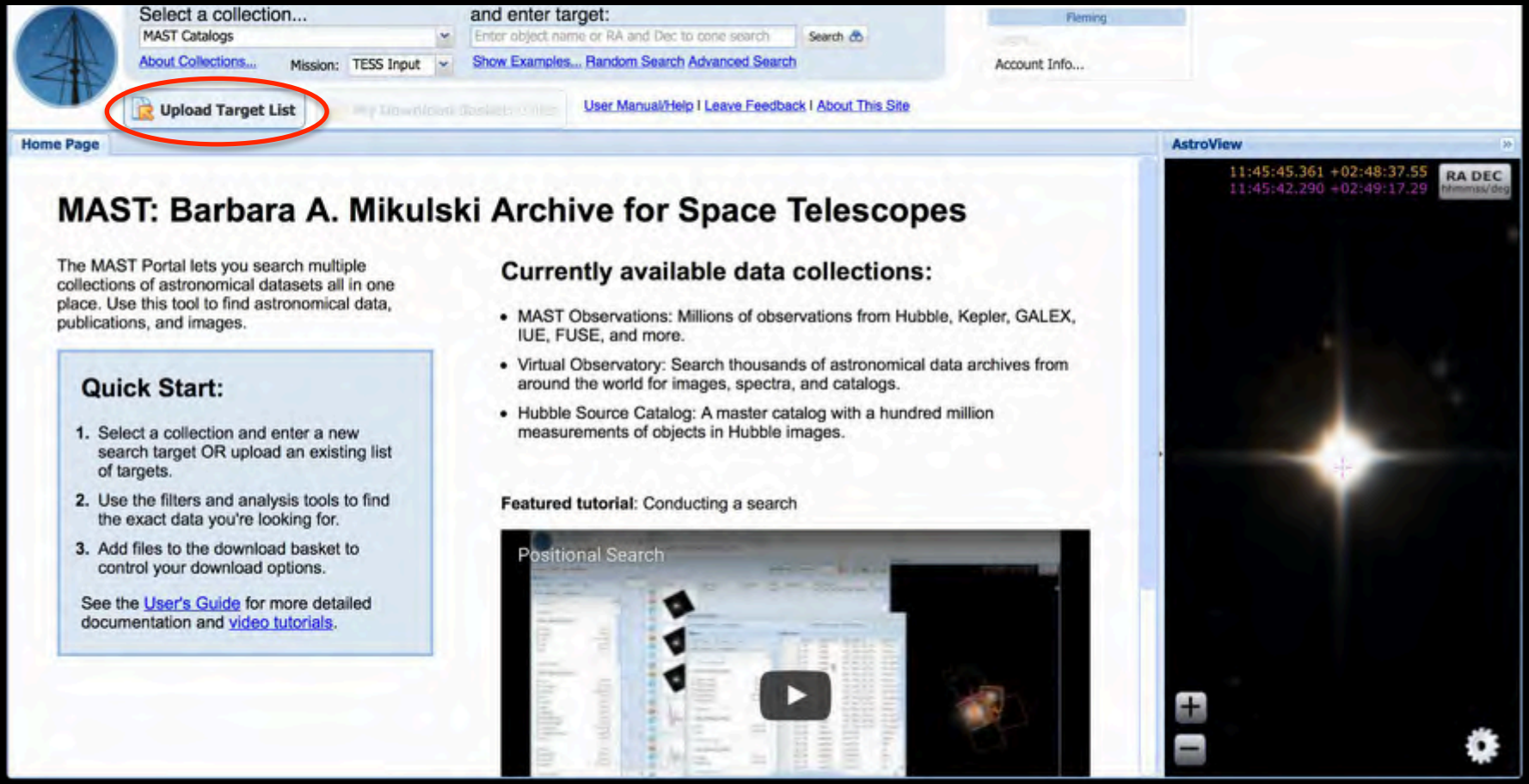
MAST Portal: Advanced Search



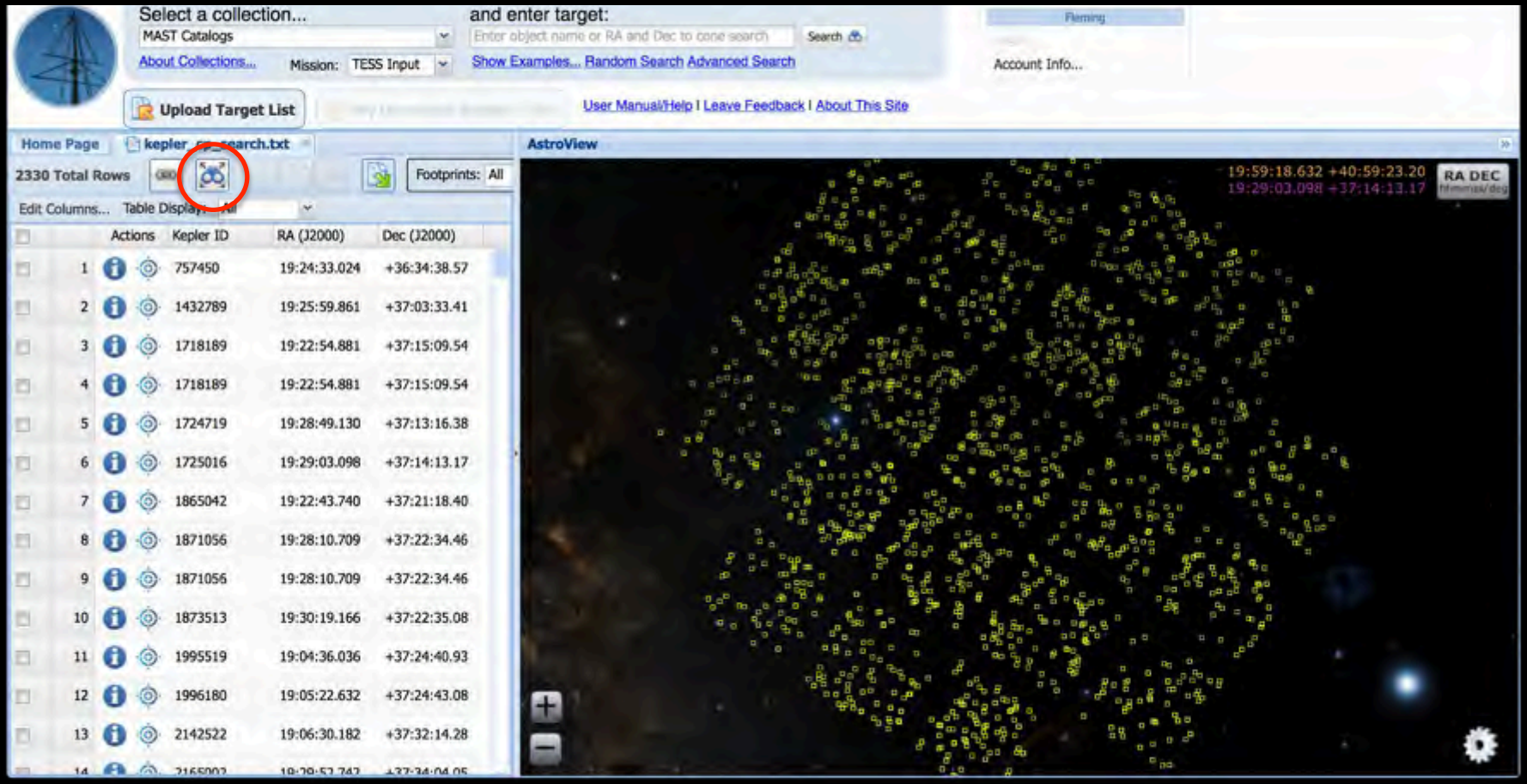
MAST Portal: Advanced Search



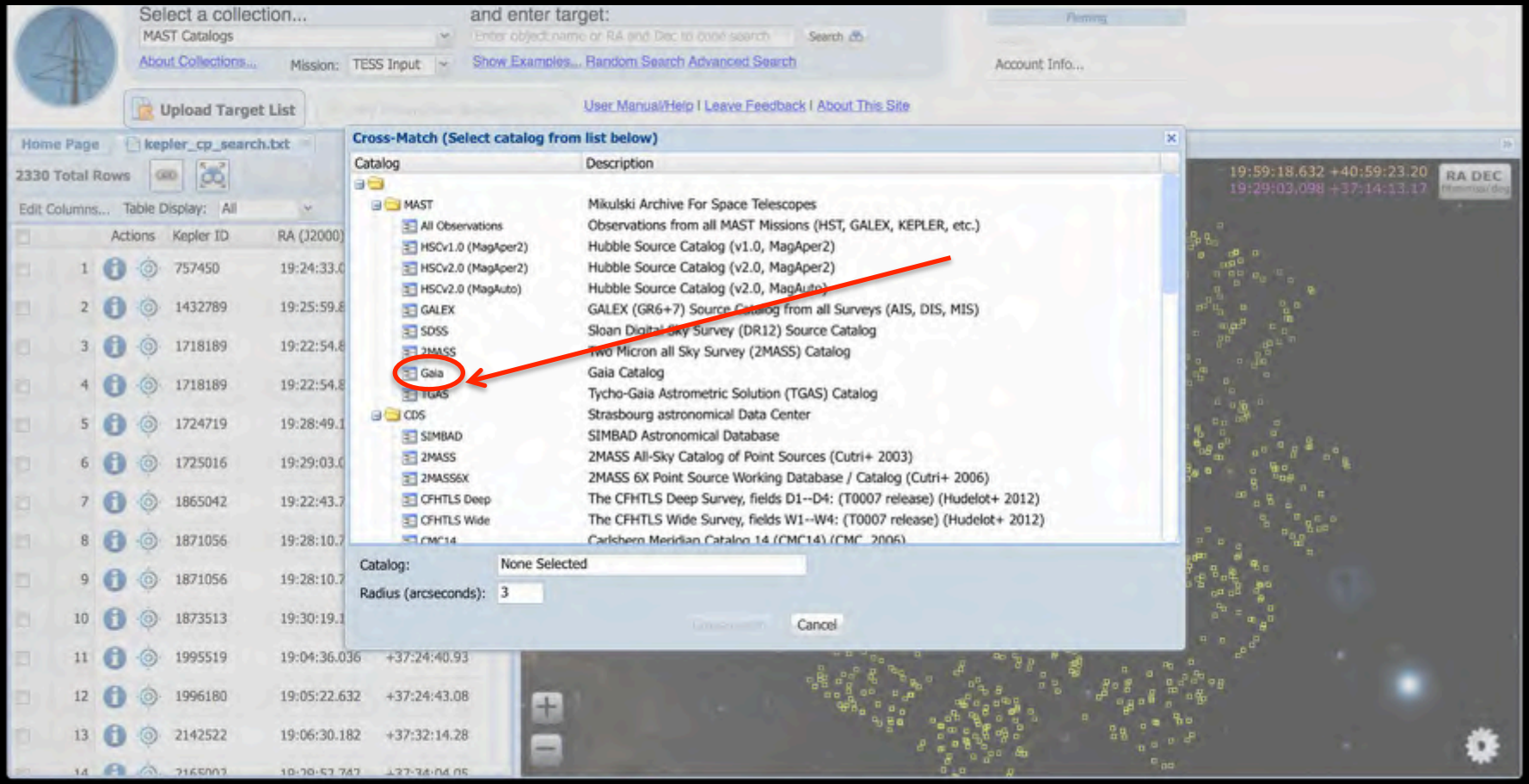
MAST Portal: Cross-Match



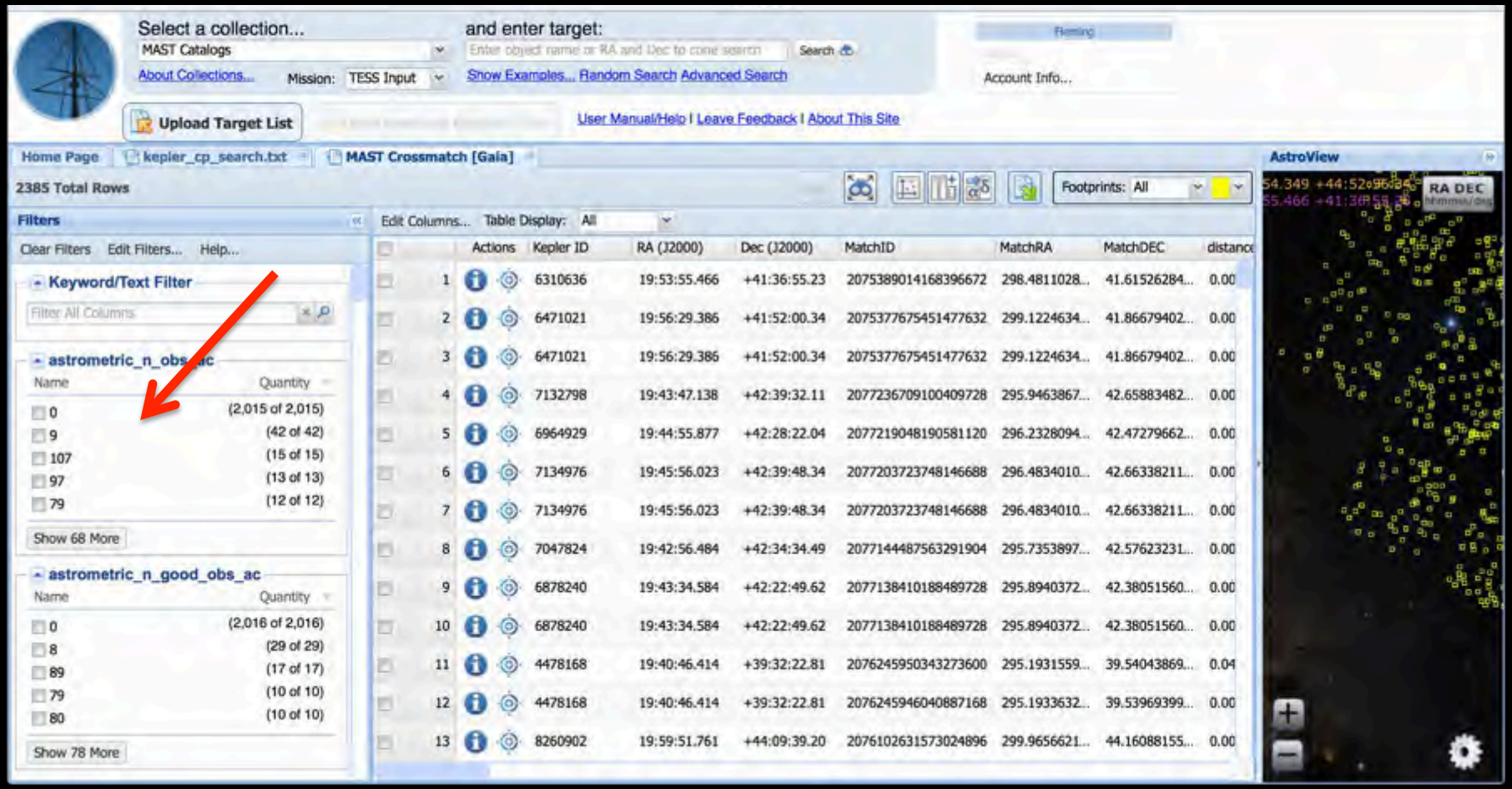
MAST Portal: Cross-Match



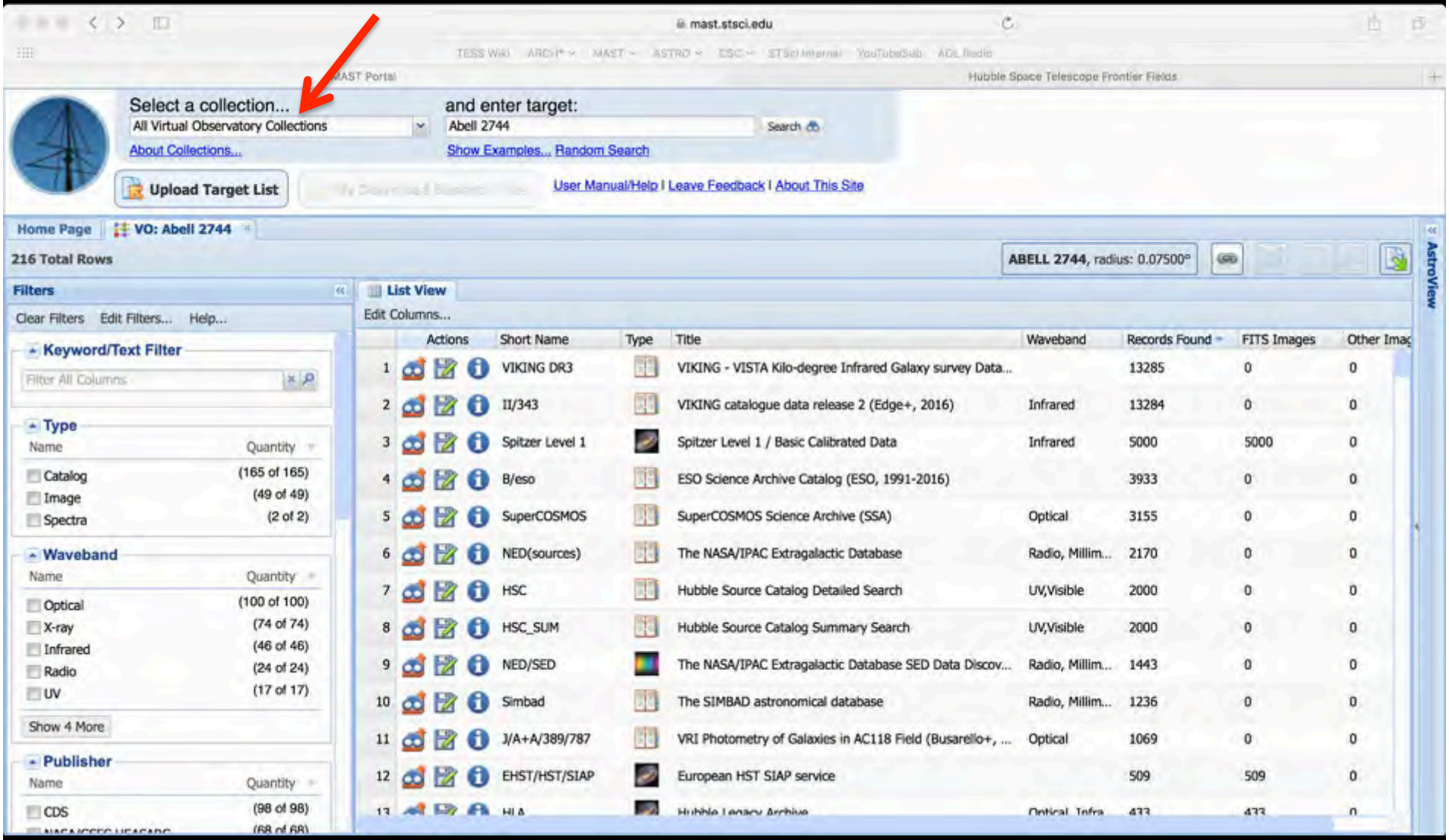
MAST Portal: Cross-Match



MAST Portal: Cross-Match



MAST Portal: Access to the Virtual Observatory



	Actions	Short Name	Type	Title	Waveband	Records Found	FITS Images	Other Imag
1		VIKING DR3		VIKING - VISTA Kilo-degree Infrared Galaxy survey Data...		13285	0	0
2		II/343		VIKING catalogue data release 2 (Edge+, 2016)	Infrared	13284	0	0
3		Spitzer Level 1		Spitzer Level 1 / Basic Calibrated Data	Infrared	5000	5000	0
4		B/eso		ESO Science Archive Catalog (ESO, 1991-2016)		3933	0	0
5		SuperCOSMOS		SuperCOSMOS Science Archive (SSA)	Optical	3155	0	0
6		NED(sources)		The NASA/IPAC Extragalactic Database	Radio, Millim...	2170	0	0
7		HSC		Hubble Source Catalog Detailed Search	UV,Visible	2000	0	0
8		HSC_SUM		Hubble Source Catalog Summary Search	UV,Visible	2000	0	0
9		NED/SED		The NASA/IPAC Extragalactic Database SED Data Discov...	Radio, Millim...	1443	0	0
10		Simbad		The SIMBAD astronomical database	Radio, Millim...	1236	0	0
11		J/A+A/389/787		VRI Photometry of Galaxies in AC118 Field (Busarello+, ...	Optical	1069	0	0
12		EHST/HST/SIAP		European HST SIAP service		509	509	0
13		HLA		Hubble Legacy Archive	Optical, Infra	433	433	0

Enhancing Data Products: gPhoton

gPhoton: A 1.1 Trillion Row Database of GALEX Photon Events

GALEX Overview

- Launched 28 April 2003
- FUV + NUV (simultaneously)
- FUV = 1350-1750 Å
- Phot. (5-6” resolution)
- Microchannel Plate Detector

Retired 28 June 2013
1.2 deg. diameter FoV
NUV = 1750-2750 Å
Spec. (R ~ 100-250)
77% of the Sky at Diff. Depths

gPhoton

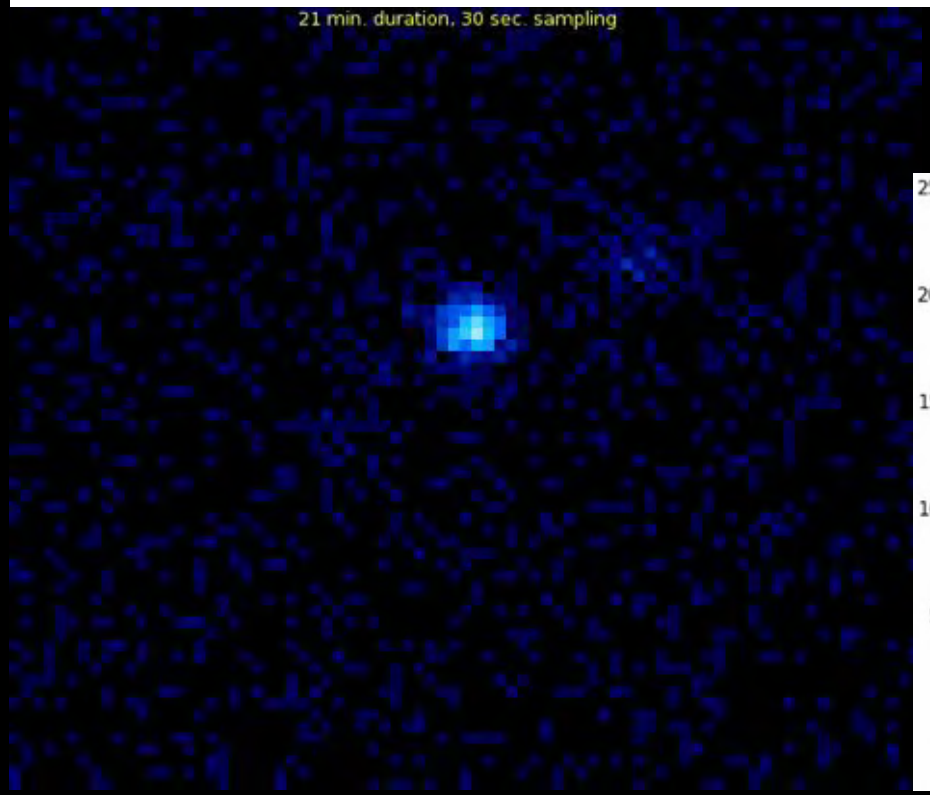
- Databse of 1.1 trillion photon events
- Open-source Python software: create cal. images and light curves
- Photon events meas. with 5 millisecond precision
- Light curves with few second sampling for any GALEX image

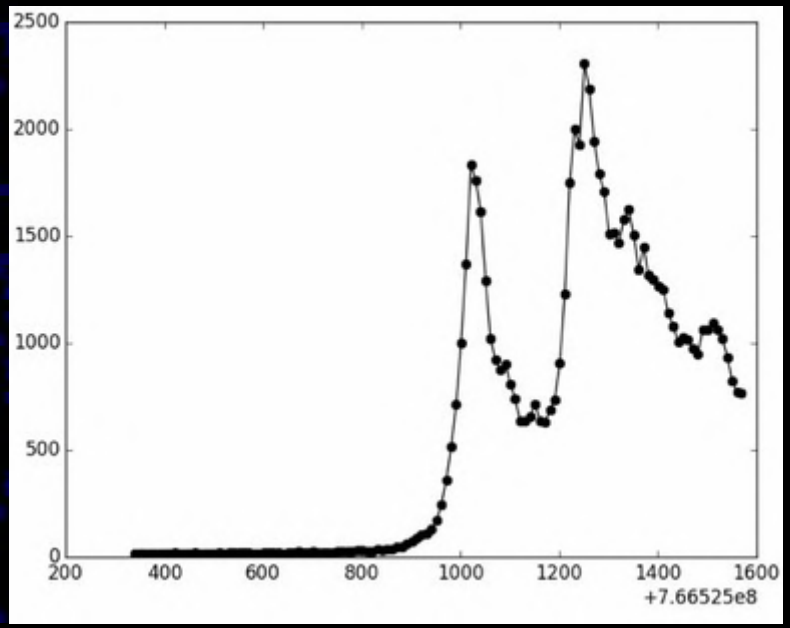
gPhoton page: https://archive.stsci.edu/prepds/gphoton/

Paper: Million, C; Fleming, S.W., Shiao, B., et al. 2016, *ApJ*, 833, 292

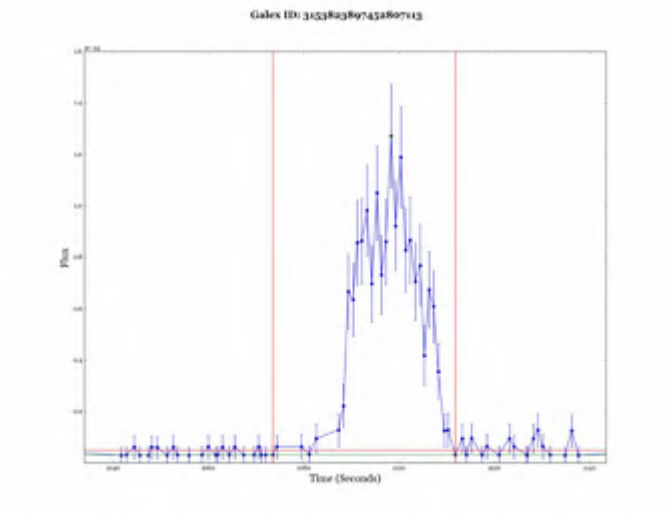
gPhoton Design Considerations

```
import gPhoton

def main():
    gPhoton.gMap(band='NUV', skypos=[176.91975, 0.25561], stepsz=30., skyrange=[0.0333,
0. 0333], cntfile='gj_3685a_movie.fits')

if __name__ == '__main__':
    main()
```

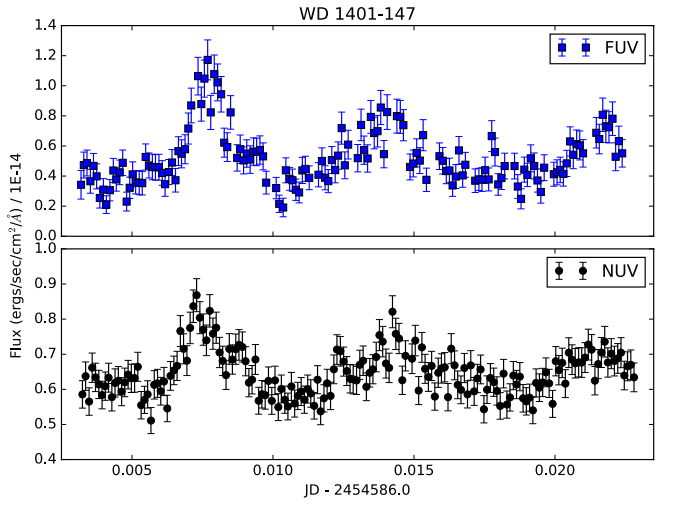




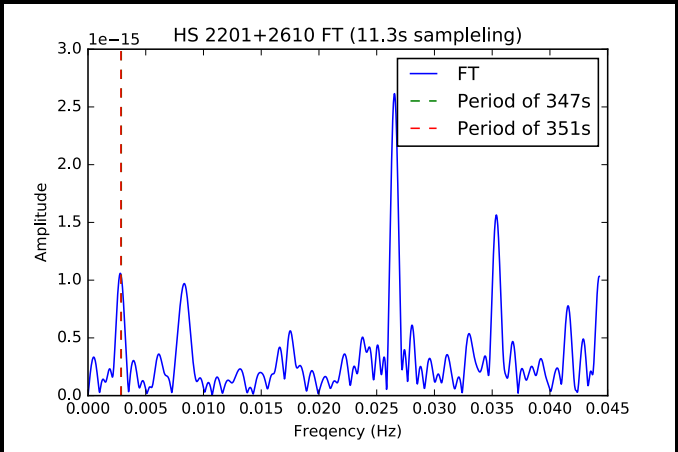
gPhoton – Early Science Examples



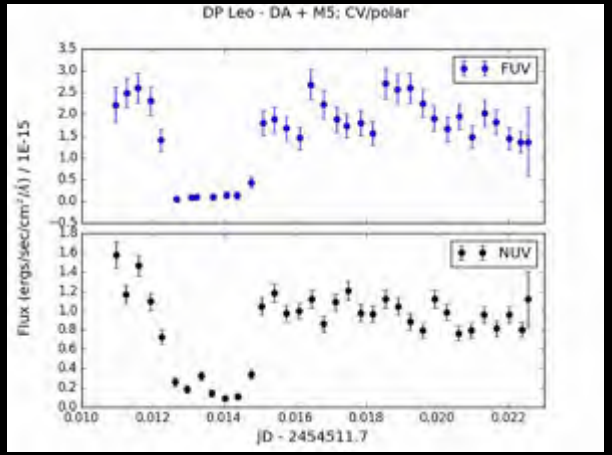
Stellar flares at low energies, w/ Rachel Osten and Clara Brasseur



White dwarf pulsations, with Michael Tucker (2015 REU)



sdB pulsation survey, with Thomas Boudreaux (2016 REU)



Eclipsing objects, including polars and WD exoplanets

Enhancing Data Products: Hubble Source Catalog

The Hubble Source Catalog (HSC)

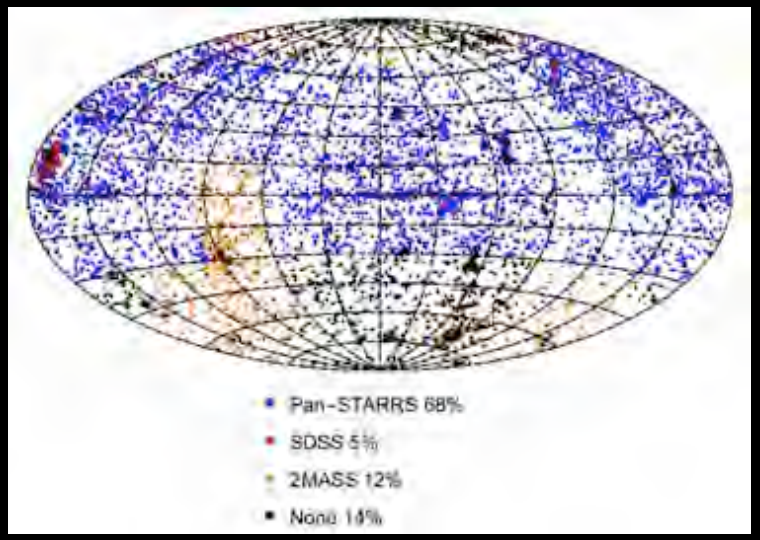
https://archive.stsci.edu/hst/hsc/

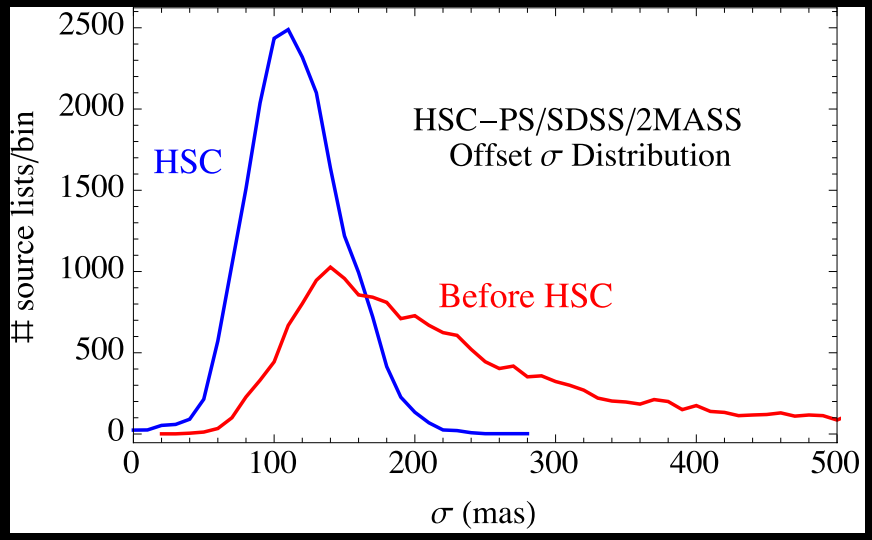
Paper: Whitmore, B.C, et al. 2016, *AJ*, 151, 134

1. Combines tens of thousands of SourceExtractor HLA source lists into a single master catalog. Uses matching algorithm from Budavari and Lubow 2012.

2. Includes WFPC2, ACS/WFC, and WFC3.

3. Absolute astrometry is good to ~100 mas (calibrated using PanSTARRS and 2MASS). This can eventually be improved to ~10 mas using Gaia observations).





HSC: Science Use Case

There are several Science Use Cases with step-by-step instructions and screen shots on the HSC webpage.

2. **Are there Use Cases available for the HSC?**

Yes. We have a variety of Use Cases:

HSC Use Case #1 - Using the Discovery Portal to Query the HSC - (Stellar Photometry in M31 - Brown et al. 2009)

HSC Use Case #2 - Using CASJOBS to Query the HSC - (Globular Clusters in M87 and a Color Magnitude Diagram for the SMC)

HSC Use Case #3 - Using the Discovery Portal to search for Variable Objects in the HSC - (Time Variability in the dwarf irregular galaxy IC 1613)

HSC Use Case #4 - Using the Discovery Portal to perform cross-matching between an input catalog and the HSC - (Search for the Supernova 2005cs progenitor in the galaxy M51)
NOTE: This use case was made using version 1. However, most of the changes are relatively minor, hence it is still quite useful.

HSC Use Case #5 - Using the Discovery Portal and CasJobs to search for Outlier Objects in the HSC - (White dwarfs in the Globular Cluster M4)

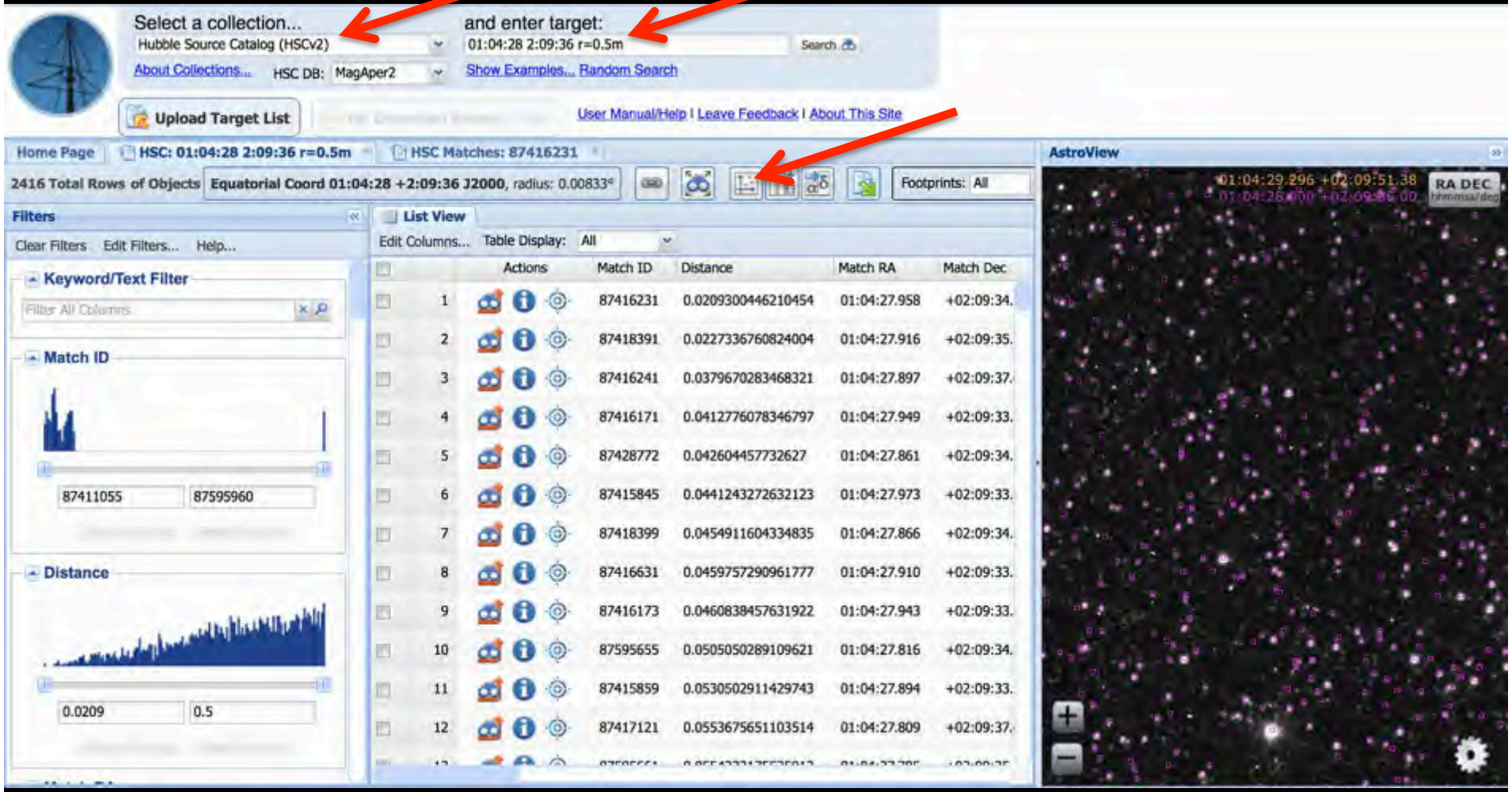
HSC Use Case #6 - Using the Discovery Portal to study the Red Sequence in a Galaxy Cluster - (The Red Sequence in the Galaxy Cluster Abell 2390)
NOTE: This use case was made using version 1. However, most of the changes are relatively minor, hence it is still quite useful.

HSC Use Case #7 - Comparing HSC "Sloan" filter magnitudes and SDSS magnitudes - (using the field around GRB110328A)<
NOTE: This use case was made using version 1. However, most of the changes are relatively minor, hence it is still quite useful.

HSC Use Case #8 - Combining HSC magnitudes and HST spectra to Investigate Objects in the HSC (using objects in the LMC Cluster R136)

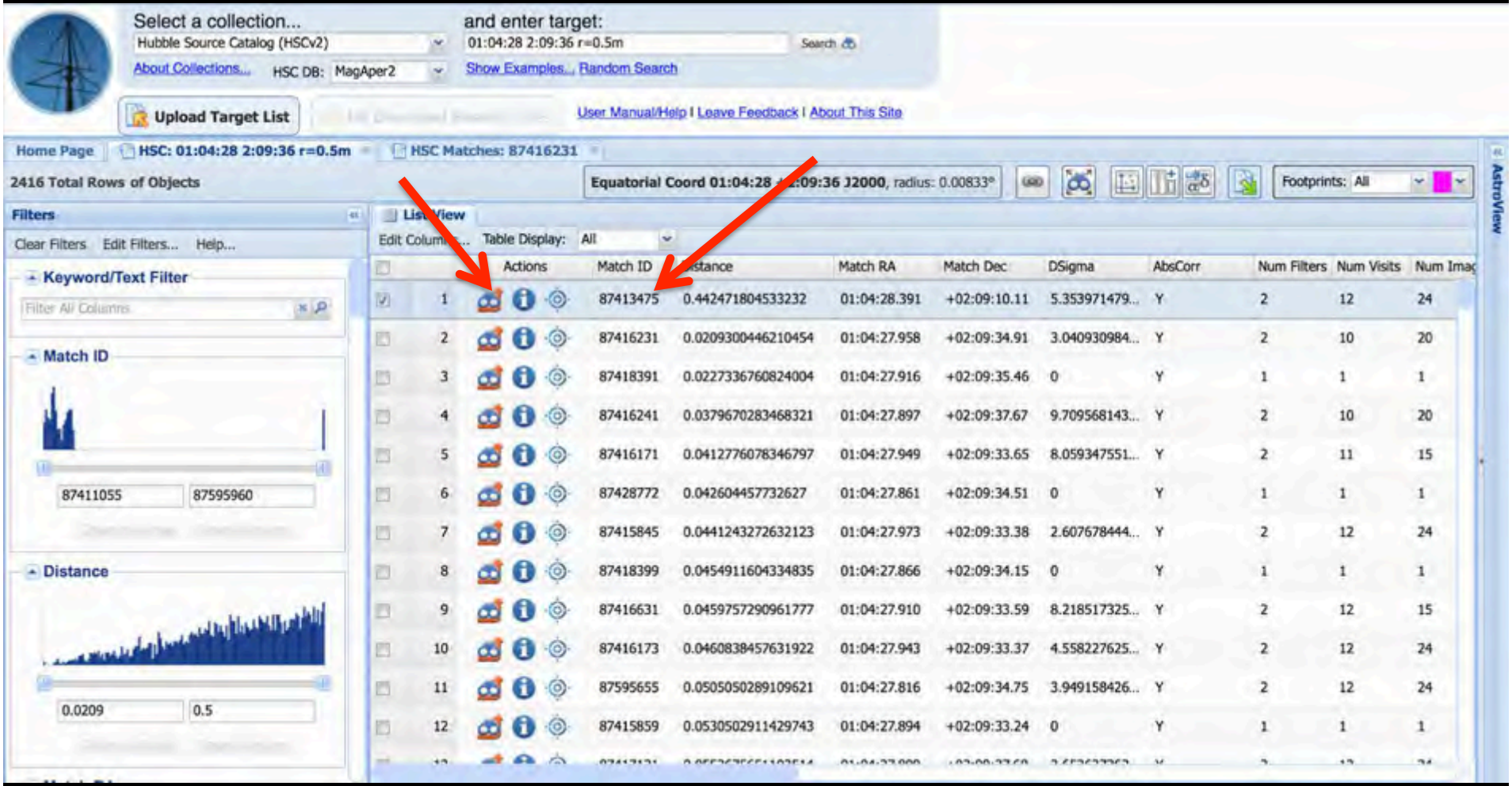
HSC Use Case #9 - Searching for Objects with both HST Imaging and Spectroscopic Data

HSC: Science Use Case, Variables in IC 1613

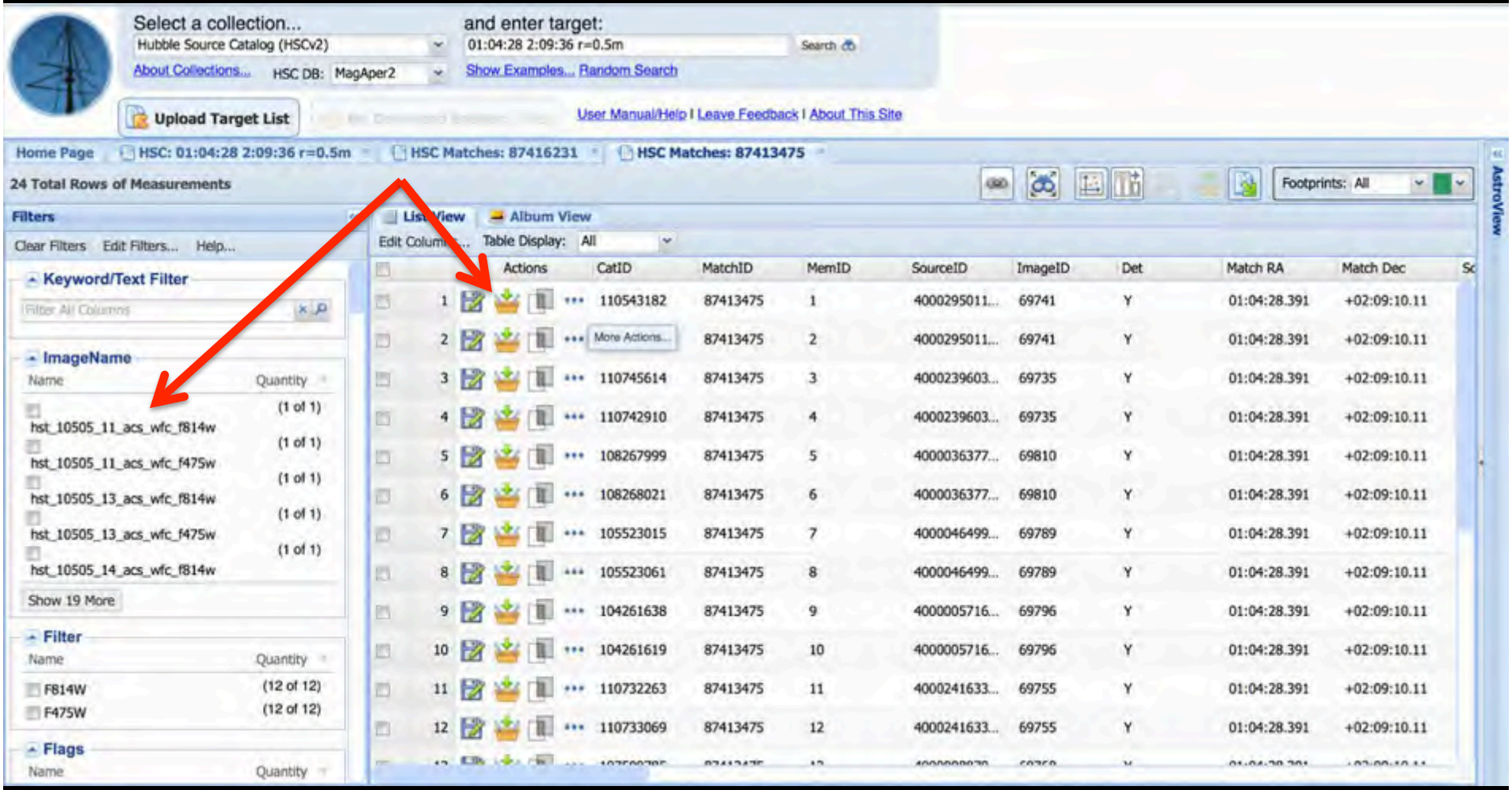


HSC: Science Use Case, Variables in IC 1613

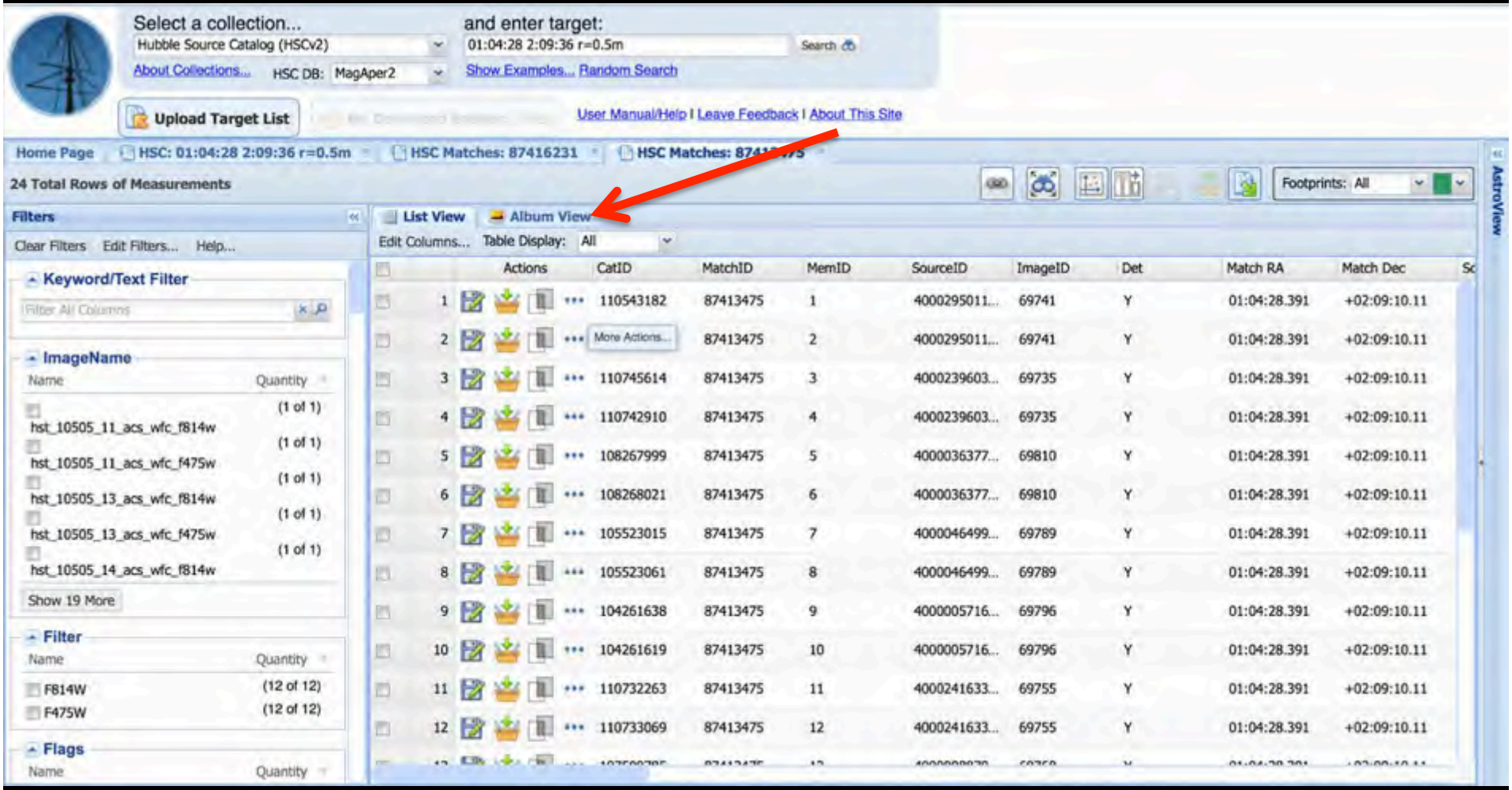
HSC: Science Use Case, Variables in IC 1613



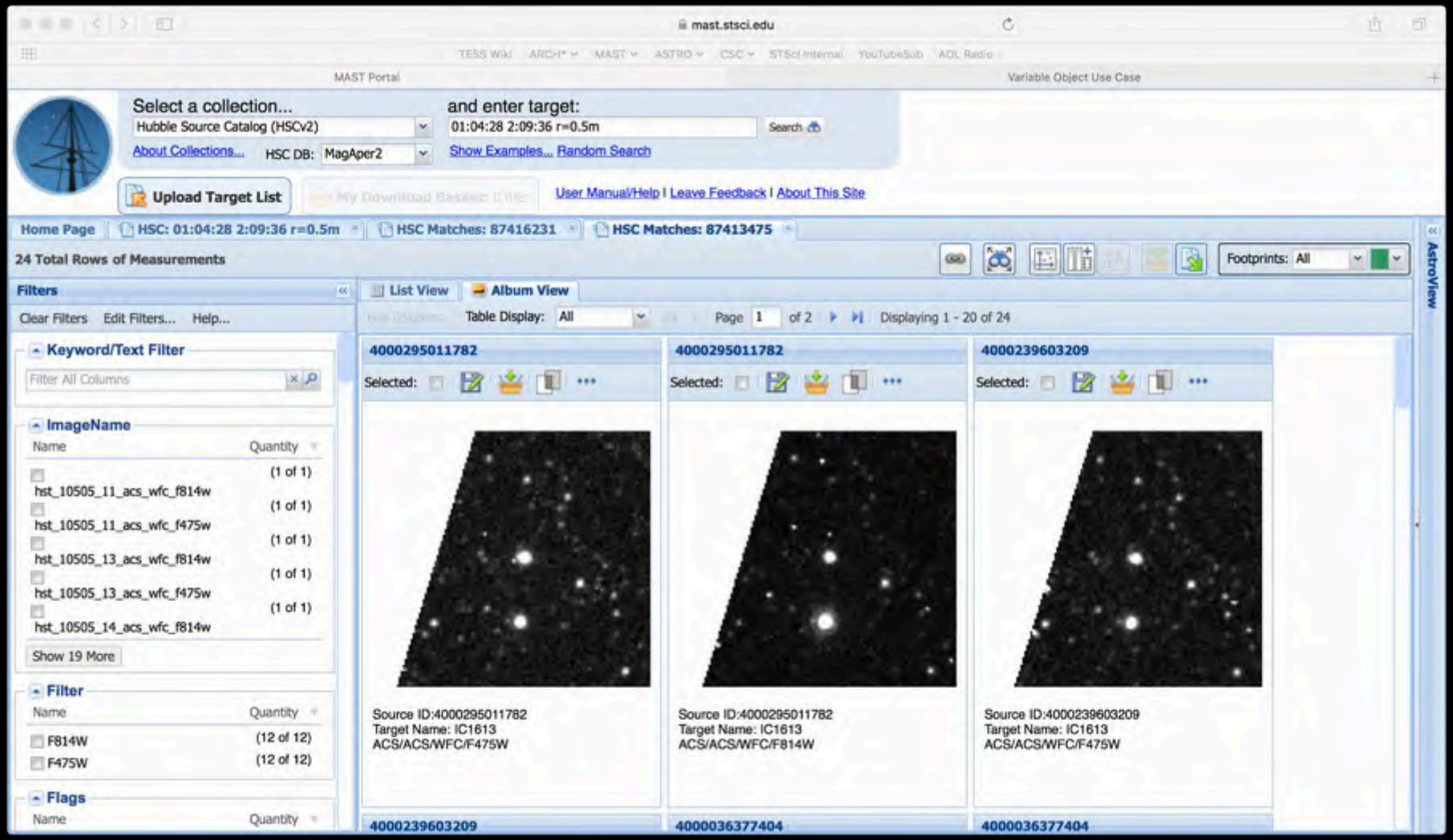
HSC: Science Use Case, Variables in IC 1613



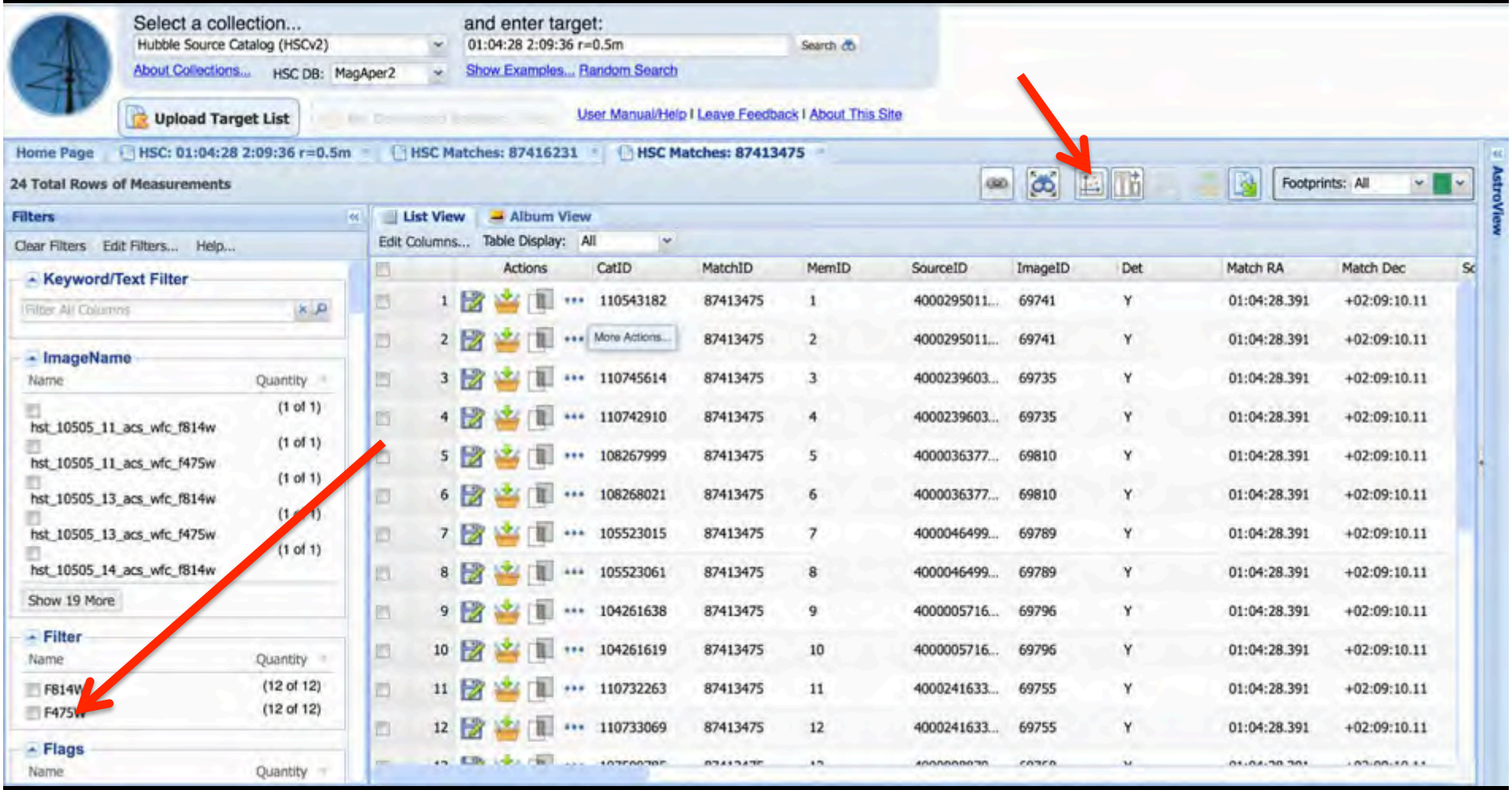
HSC: Science Use Case, Variables in IC 1613



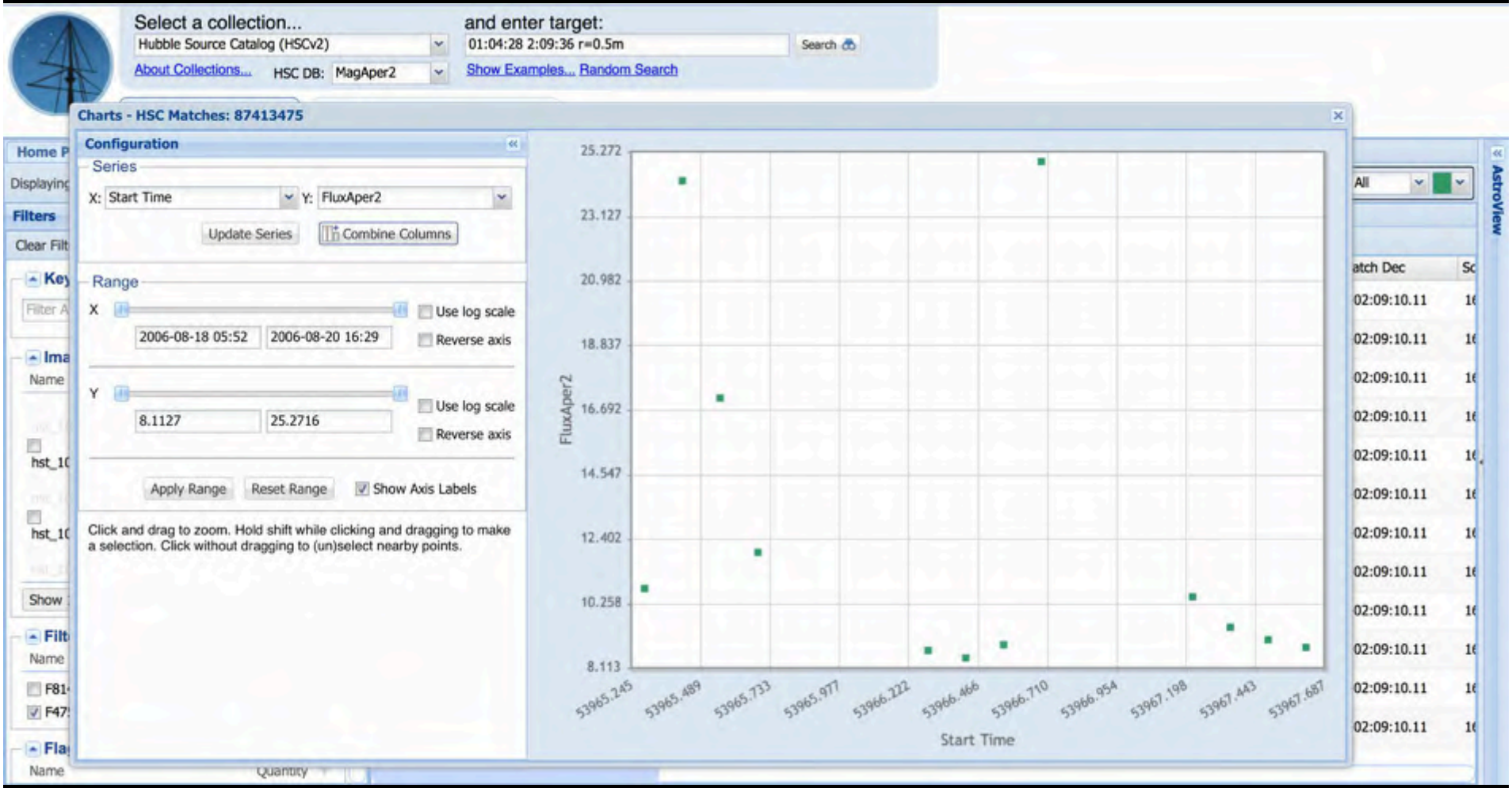
HSC: Science Use Case, Variables in IC 1613



HSC: Science Use Case, Variables in IC 1613



HSC: Science Use Case, Variables in IC 1613



Summary

- MAST supports more than 20 missions: cross-mission discoverability is a key driver in our development efforts.
- The MAST Portal allows for searching across MAST missions and catalogs, the Virtual Observatory, and CDS catalogs.
- Beyond archving mission data, MAST also creates enhanced data products that enable new science. Recent examples are:
  - gPhoton: A time-tagged databse of 1.1 trillion GALEX photon events.
  - HSC: A master catalog of unique sources across all HST observations.
- Newly created STScI Data Mission Office, led by Arfon Smith, will begin to expand MAST capabilities into modern technological arenas, including cloud storage, “bring-code-to-the-data” environments, improved access to high performance computing resources, better support for API access to MAST resources.

The NOAO Data Lab Project
Introduction

Knut Olsen

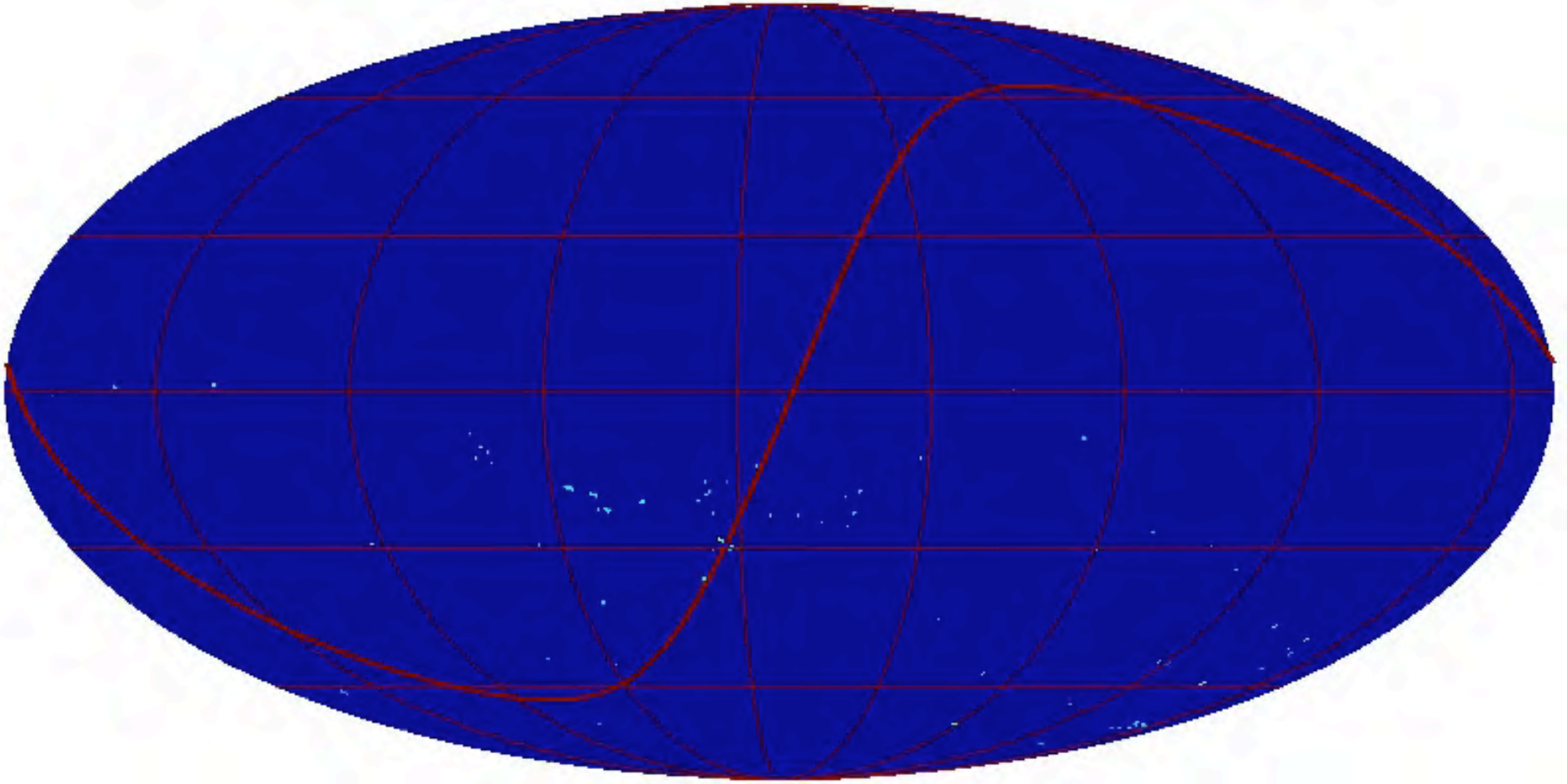
for the Data Lab team

Data Lab Team

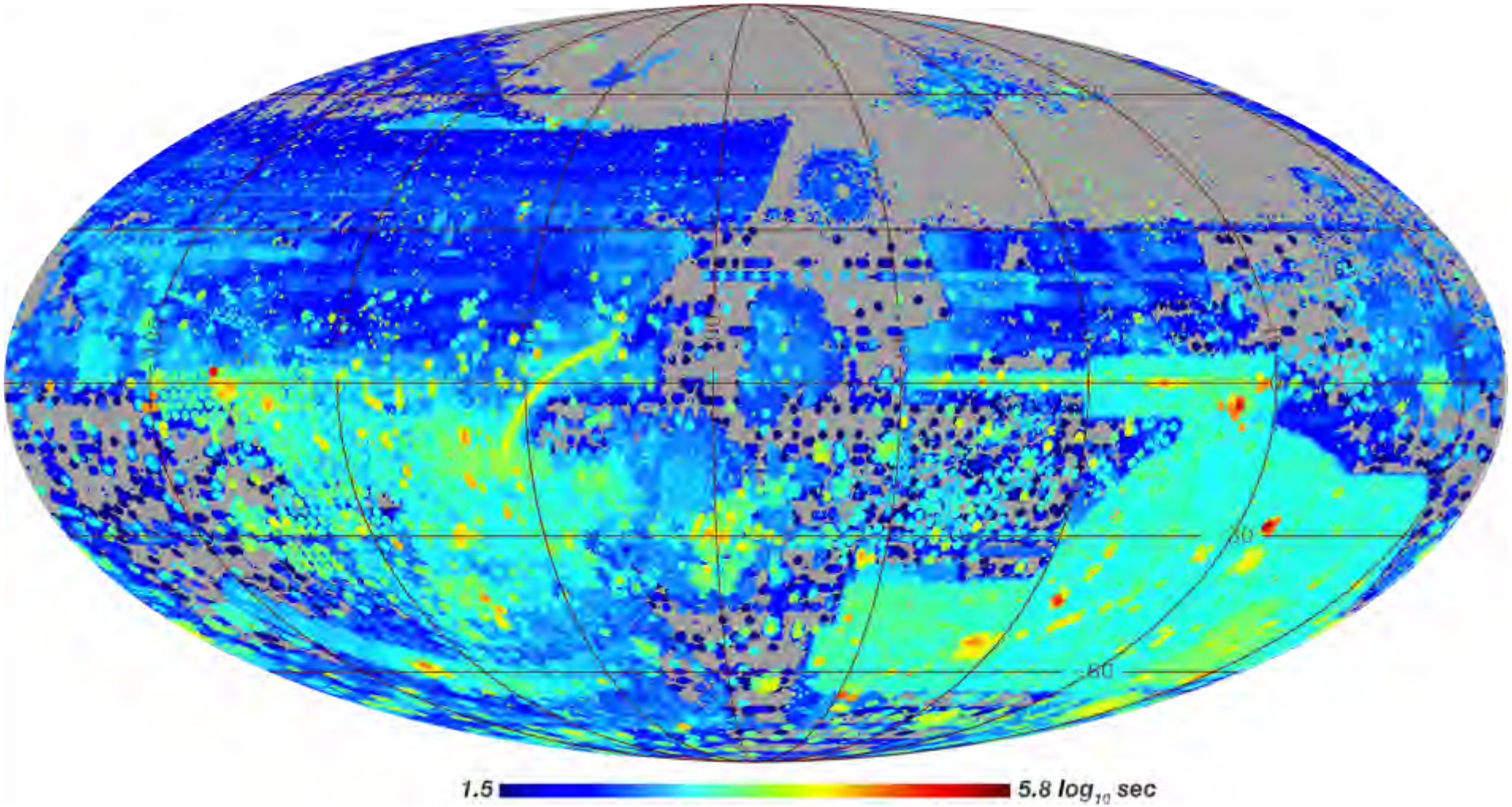
Current team:

- Mike Fitzpatrick, Lead Developer
- Matthew Graham, Scientist/Developer
- Wendy Huang, Software Engineer
- Stephanie Juneau, Data Scientist
- David Nidever, Data Scientist
- Robert Nikutta, Data Scientist
- Pat Norris, Test Engineer
- Knut Olsen, Project Scientist
- Steve Ridgway, Scientist
- Pete Wargo, System Administrator

NOAO wide field imaging data over time



DECam and Mosaic data in May 2016



Data Volume and Complexity

500 TB (January 2017) of on-target imaging data (t_{exp}>30s) currently from:

- Dark Energy Survey
- Legacy Surveys for DESI Targeting
- Community DECam and Mosaic programs and surveys

Hundreds of TB more coming

Total holdings of several PB

Large catalogs coming:

- Dark Energy Survey – 45 TB
- DESI Targeting Survey – ~5 TB
- Community programs and surveys – up to several TB each

NOAO Data Lab

- Goal:
- Efficient exploration and analysis of the large datasets being generated by instruments on NOAO wide-field 4-m telescopes
- Approach:
  - Catalogs and images linked to catalog objects
  - Data discovery
  - Developing intuition through interaction with selected catalog and image set of known objects
  - Automation of analysis to aid discovery of unknown objects

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

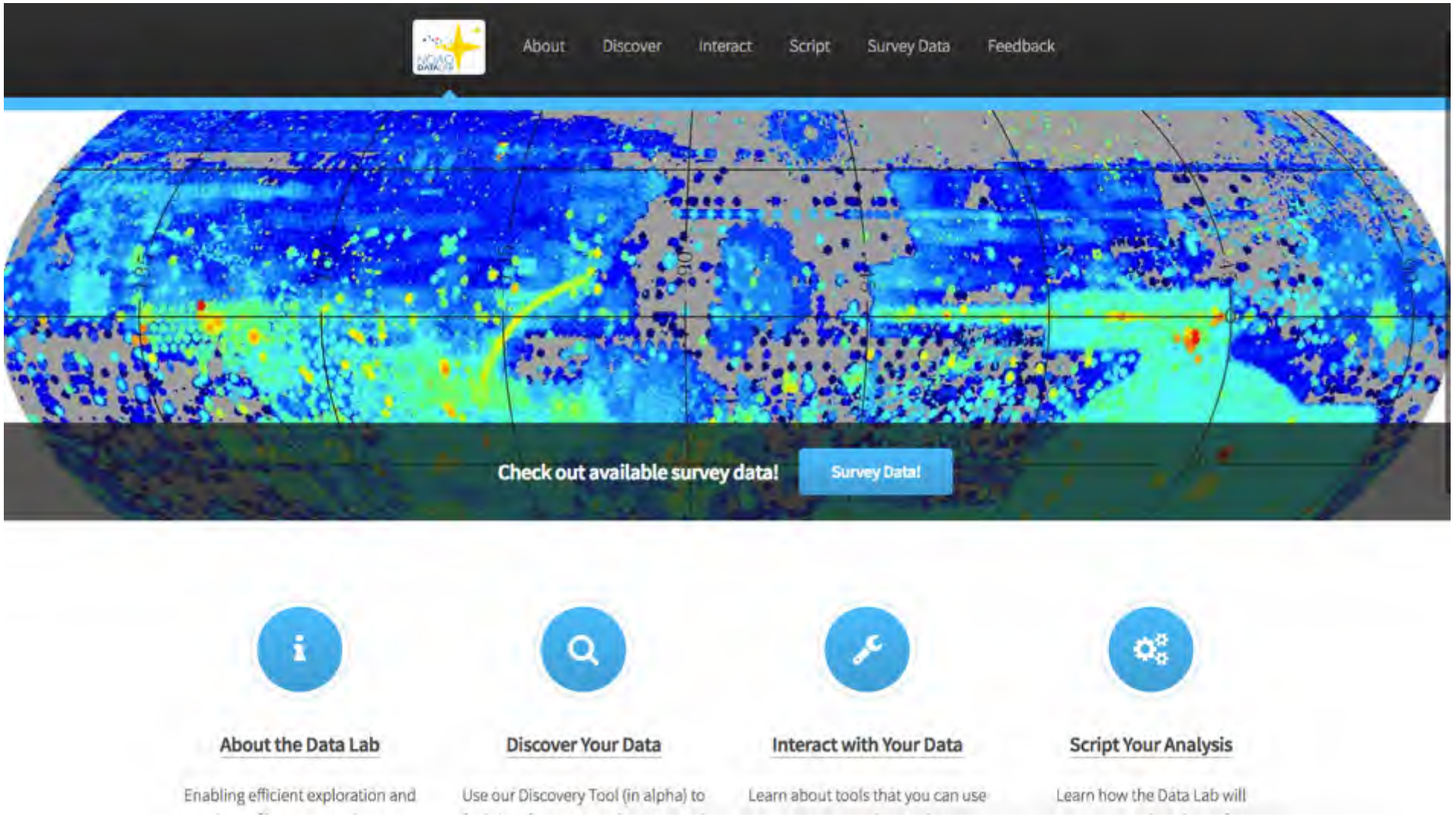
Compute Processing – Data Lab will allow workflows to run close to the data

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

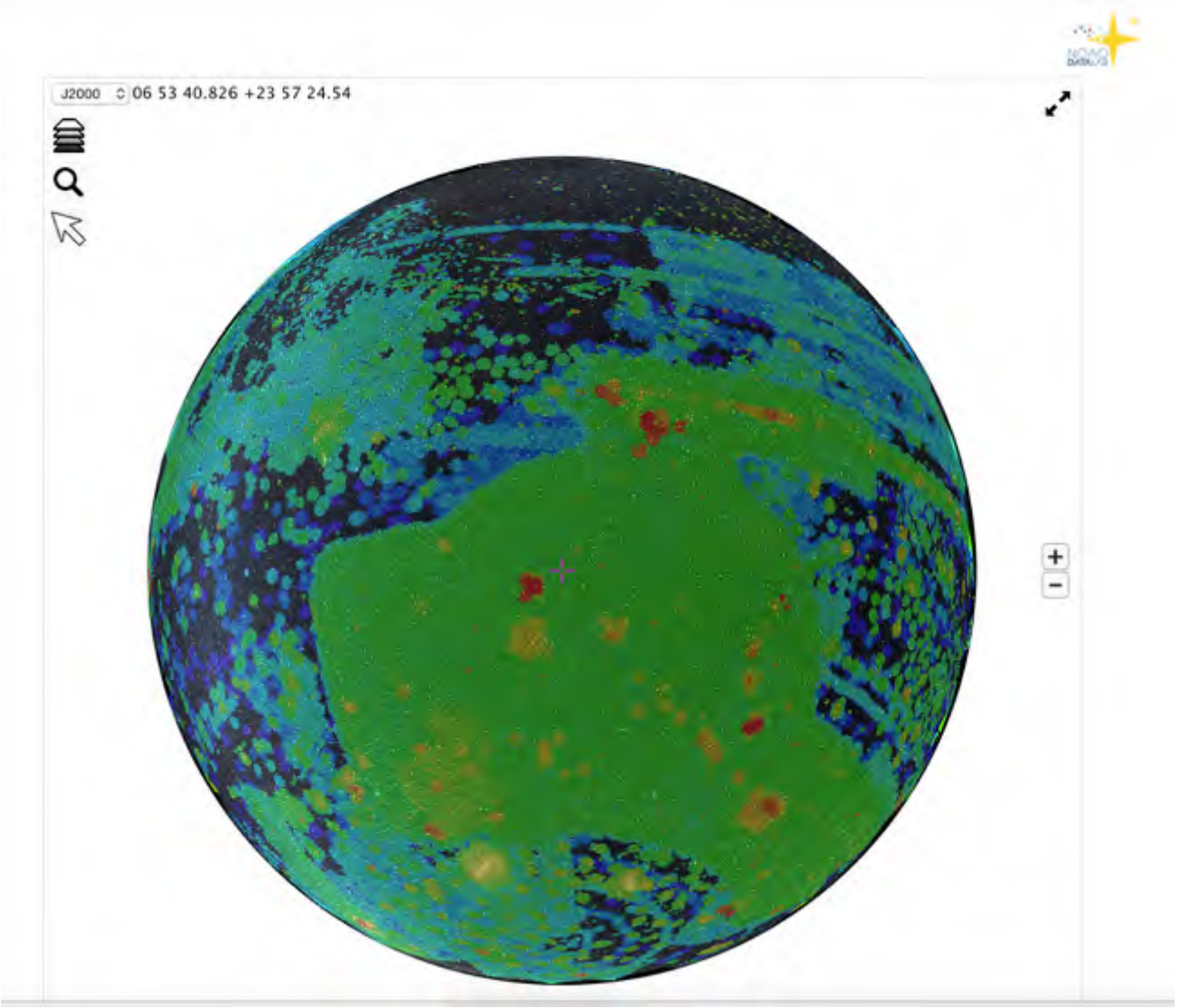
Timeline

- March 2015: Conceptual Design Review
  - Lisa Storrie-Lombardi (Chair), Severin Gaudet, Zeljko Ivezic, Connie Rockosi, Beth Willman reviewed Science Case & Requirements, System Architecture, Operations Concept & Requirements, and Schedule
- Fall 2015 hiring campaign
- June 2016 San Diego AAS Demo
- August 2016 Interim Review
  - Lisa Storrie-Lombardi (Chair), Severin Gaudet, Zeljko Ivezic, Ed Olszewski, Beth Willman, and Dennis Zaritsky reviewed progress and Year 2 plan
- January 2017 AAS SMASH DR1 and DECaLS DR3
- Summer 2017 first public release
- End 2017/Early 2018 DES DR1

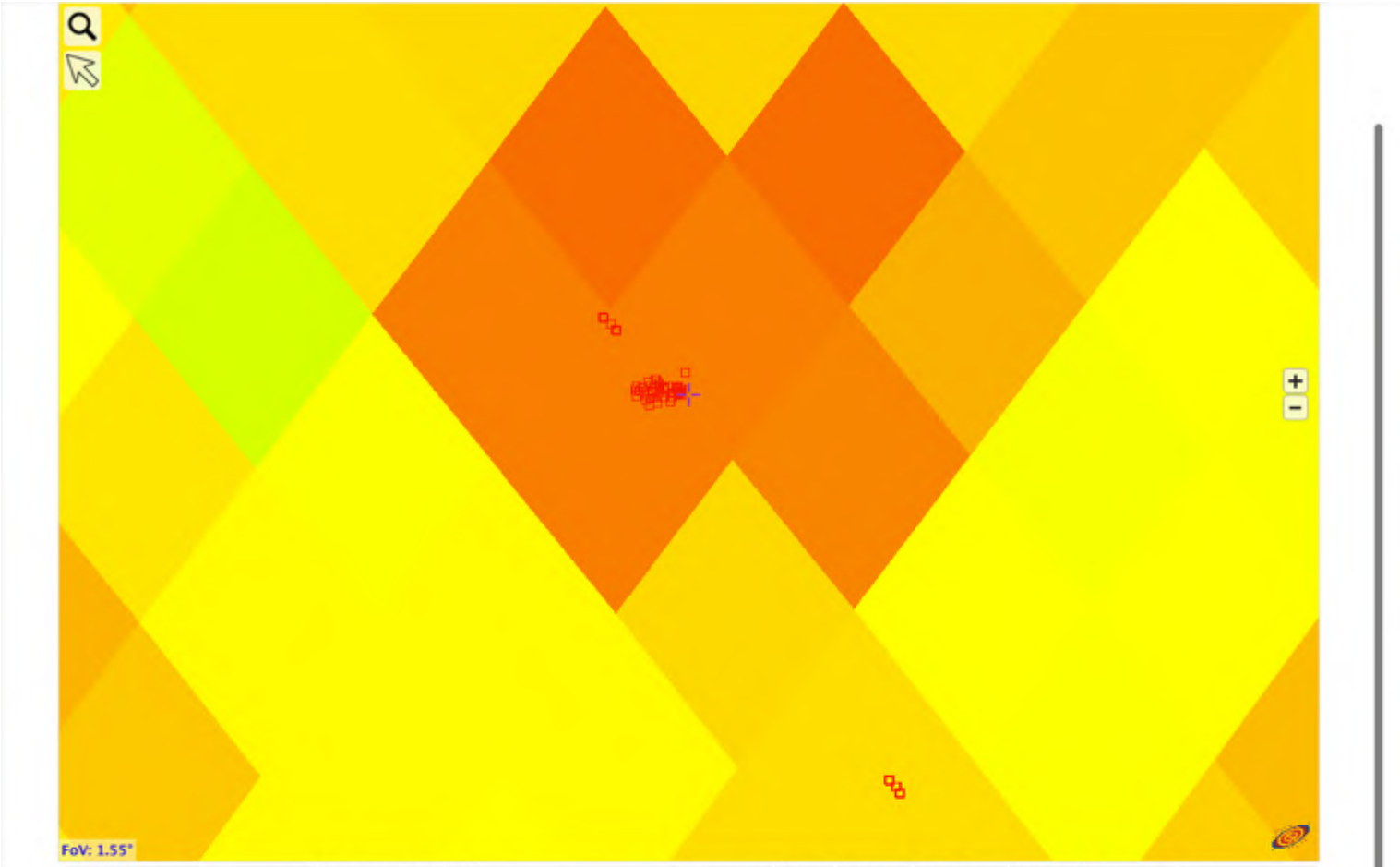
datalab.noao.edu



Data discovery

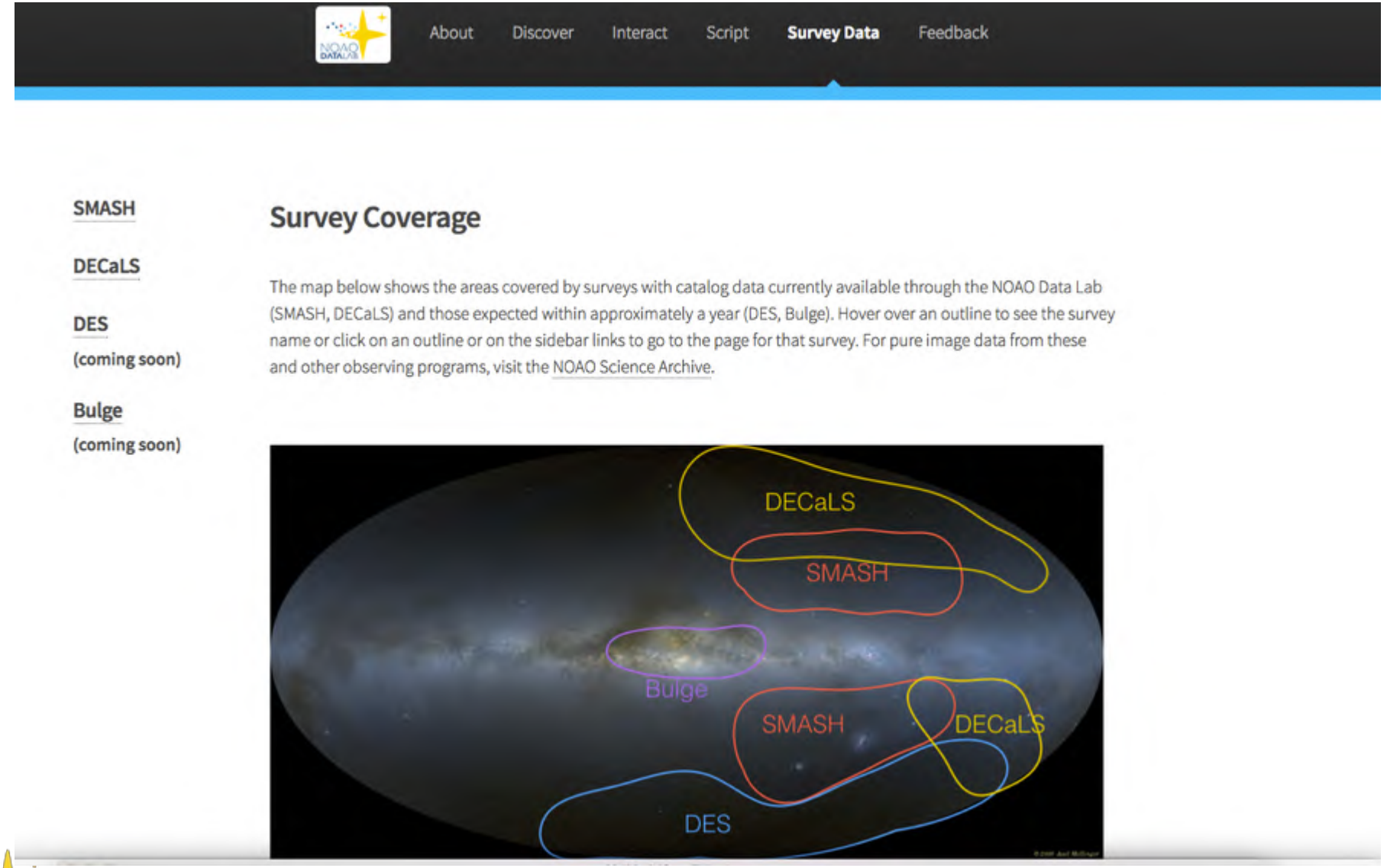


Data discovery



REFERENCE	FITS_EXTENSION	OBJECT	SURVEY	SURVEYID	PROP_ID	START_DATE	RA
c4d_151112_003815_ori.fits.fz	1	NGC1399_t1_d3_short			2015B-0602	2015-11-11 00:00:00	54.5602958333333
c4d_151112_020032_ori.fits.fz	1	NGC1399_t1_d3_short			2015B-0602	2015-11-11 00:00:00	54.5602041666667
c4d_151111_020907_ori.fits.fz	1	NGC1399_t1_d11_short			2015B-0602	2015-11-10 00:00:00	54.56775
c4d_151111_074627_ori.fits.fz	1	NGC1399_t1_d13_short			2015B-0602	2015-11-10 00:00:00	54.5581208333333
c4d_151112_001200_ori.fits.fz	1	NGC1399_t1_d1_short			2015B-0602	2015-11-11 00:00:00	54.5554208333333
c4d_151112_011749_ori.fits.fz	1	NGC1399_t1_d1_short			2015B-0602	2015-11-11 00:00:00	54.5571291666667
							54.5552541666667

Survey data


About
Discover
Interact
Script
Survey Data
Feedback
SMASH
DECaLS
DES
(coming soon)
Bulge
(coming soon)
Survey Coverage
The map below shows the areas covered by surveys with catalog data currently available through the NOAO Data Lab (SMASH, DECaLS) and those expected within approximately a year (DES, Bulge). Hover over an outline to see the survey name or click on an outline or on the sidebar links to go to the page for that survey. For pure image data from these and other observing programs, visit the NOAO Science Archive.
DECaLS
SMASH
Bulge
SMASH
DECaLS
DES

SMASH DR1

Survey of the Magellanic Stellar History (SMASH)

About Discover Interact Script **Survey Data** Feedback

Description

- Overview
- Goals
- First Data Release
- Data Reduction and Calibration

Data Access

Analysis

Explore

Results

The SMASH Survey

Overview

The Survey of the Magellanic Stellar History (SMASH) is using DECam to map 480 square degrees of sky to depths of ugriz~24 with the goal of identifying broadly distributed, low surface brightness stellar populations associated with the stellar halos and tidal debris of the Magellanic Clouds. It will eventually contain measurements of approximately 250 million objects distributed in discrete fields spanning an area of about 2400 square degrees. The first data release (DR1) contains ~100 million objects from 61 observed fields. Browse these pages to learn more about SMASH and to access the data. The SMASH overview paper (Nidever et al. 2017) describes the survey in detail, including its goals, survey strategy, reduction, and calibration.

SMASH DR1 Data Access

Description

Data Access

Analysis

Explore

Results

Data Access

The SMASH data are accessible by a variety of means:

Data Lab Table Access Protocol (TAP) service

TAP provides a convenient access layer to the SMASH catalog database. TAP-aware clients (such as TOPCAT) can point to *http://datalab.noao.edu/tap*, select the *smash_dr1* database, and see the database tables and descriptions. *smash_dr1* contains six tables: *chip, exposure, field, object, source,* and *xmatch*. These are described in the schema page.

Data Lab Query Manager

The Query Manager is available as part of the prototype Data Lab software distribution. The Query Manager client provides a Python API to Data Lab database services. For the SMASH DR1 release, these services include only anonymous access through synchronous queries of the catalog made directly to the database. The full public release of the Data Lab Query Manager in the summer of 2017 will include authenticated access, synchronous and asynchronous queries, TAP queries, personal database storage, and storage through the Data Lab VOSpace.

Image cutouts

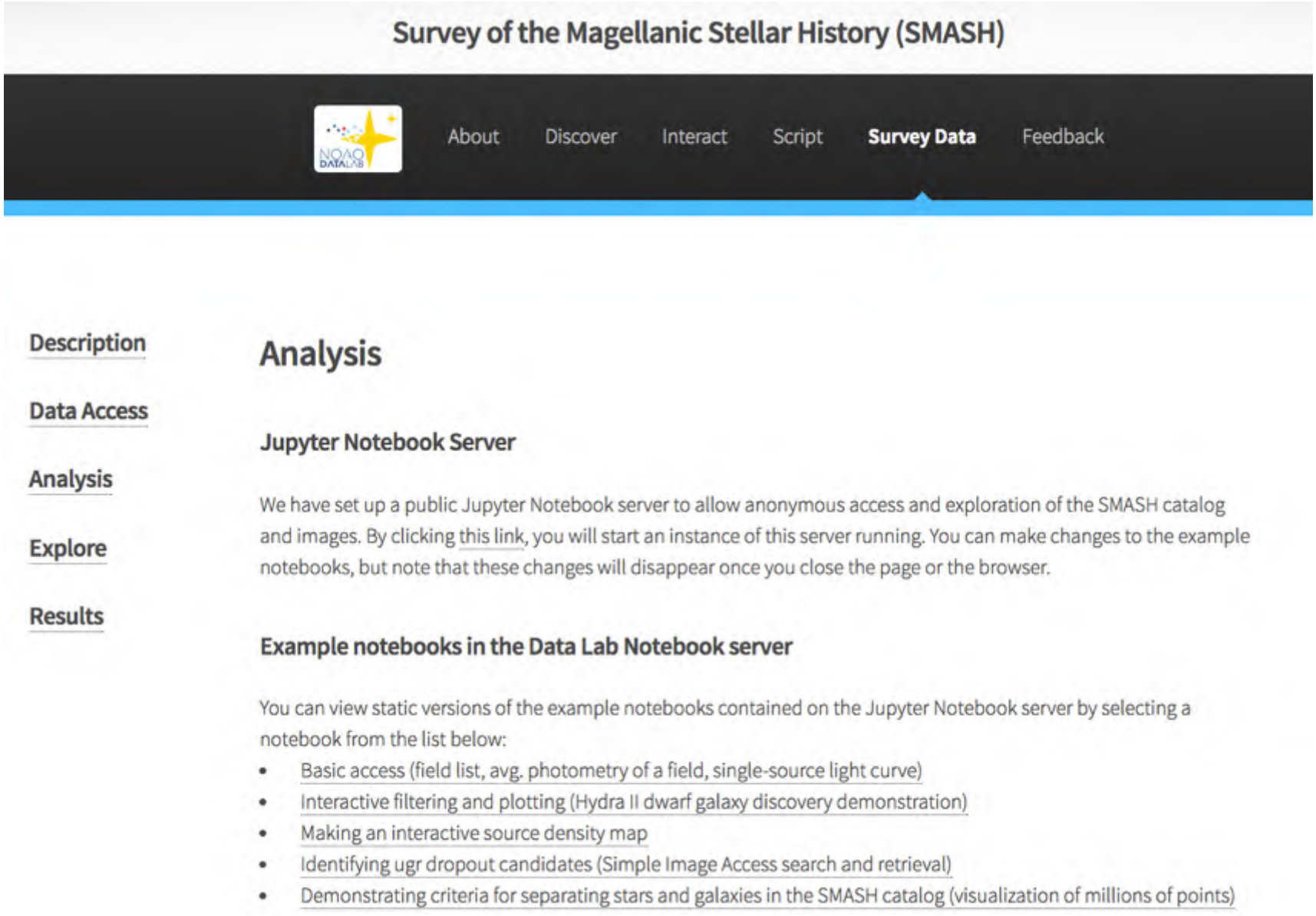
The Data Lab Simple Image Access (SIA) service provides a fast way to retrieve cutouts from SMASH images. For an example of how to use the SIA service, see this Jupyter notebook.

Jupyter Notebook Server

The Data Lab Jupyter Notebook server contains examples of how to access and visualize the SMASH catalog.

FTP access

SMASH DR1 Data Analysis



Identifying *r*-dropouts

Catalog query

For our query, we will look for objects that are undetected or have large errors in u, g, and r, but are detected and have small errors in i and z. We will only keep objects that have a match in the ALLWISE catalog. Using subqueries to limit the object and xmatch tables using indexed columns makes the query run much faster than it would otherwise.

In [2]:
```
%%time
db1='smash_dr1.object' # the SMASH object table with average magnitudes
db1sel='db1.fieldid,db1.id,db1.ra,db1.dec,db1.umag,db1.gmag,db1.rmag,db1.imag,' +\
    'db1.zmag,db1.uerr,db1.gerr,db1.rerr,db1.ierr,db1.zerr,db1.depthflag' # select ID, coordinates, and mags
db2='smash_dr1.xmatch' # the SMASH cross-match table, which contains cross-matches to ALLWISE
db2sel='db2.wise_id,db2.wise_w1mag,db2.wise_w1err,db2.wise_w2mag,db2.wise_w2err' # ALLWISE W1&W2 mags
db1where='(db1.ndetu=0 or db1.uerr>0.3) and ' + \
         ' (db1.ndetg=0 or db1.gerr>0.3) and ' + \
         ' (db1.ndetr=0 or db1.rerr>0.3) and ' + \
         ' (db1.ndeti>0 and db1.ierr<0.1) and ' + \
         ' (db1.ndetz>0 and db1.zerr<0.1)' # pick ugr dropouts
db2where='(db1.id=db2.id)' # only pick dropouts that are found in ALLWISE W1

# Create the query string.
query = 'SELECT '+db1sel+','+db2sel+' FROM (SELECT * FROM '+db1+' WHERE depthflag > 1) AS db1, '+ \
    '(SELECT * FROM '+db2+' WHERE wise_match=1) AS db2 ' +\
    'WHERE ('+db2where+' and '+db1where+')'

print "Your query is:", query
print "Making query"

# Call the Query Manager Service
response = queryClient.query(token, adql = query, fmt = 'csv')
df = pd.read_csv(StringIO(response))

print len(df), "objects found."
```

```
Your query is: SELECT db1.fieldid,db1.id,db1.ra,db1.dec,db1.umag,db1.gmag,db1.rmag,db1.imag,db1.zmag,db1.uerr,db1.ger
r,db1.rerr,db1.ierr,db1.zerr,db1.depthflag,db2.wise_id,db2.wise_w1mag,db2.wise_w1err,db2.wise_w2mag,db2.wise_w2err FR
OM (SELECT * FROM smash_dr1.object WHERE depthflag > 1) AS db1, (SELECT * FROM smash_dr1.xmatch WHERE wise_match=1) A
S db2 WHERE ((db1.id=db2.id) and (db1.ndetu=0 or db1.uerr>0.3) and  (db1.ndetg=0 or db1.gerr>0.3) and  (db1.ndetr=0 o
r db1.rerr>0.3) and  (db1.ndeti>0 and db1.ierr<0.1) and  (db1.ndetz>0 and db1.zerr<0.1))
Making query
5769 objects found.
CPU times: user 52.3 ms, sys: 11.1 ms, total: 63.3 ms
Wall time: 39.2 s
```

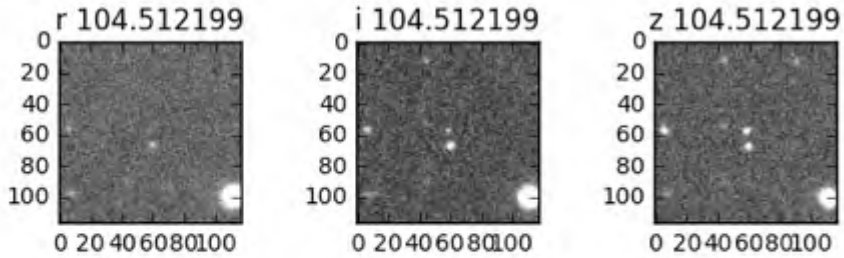
Identifying *r*-dropouts

Displaying the cutouts

Now let's show the cutouts. The object in question is indeed invisible in the r-band image, but is visible in both i and z, and appears point-like.

In [12]:
```
a1=plt.subplot2grid((2,8),(0,0),rowspan=2,colspan=2)
imgplot = plt.imshow(rimg)
a1.set_title('r '+id1.astype('string'))

a2=plt.subplot2grid((2,8),(0,3),rowspan=2,colspan=2)
imgplot = plt.imshow(iimg)
a2.set_title('i '+id1.astype('string'))

a3=plt.subplot2grid((2,8),(0,6),rowspan=2,colspan=2)
imgplot = plt.imshow(zimg)
a3.set_title('z '+id1.astype('string'))
```

Out[12]:
```
<matplotlib.text.Text at 0x7fc06272fe90>
```



To go through the whole list of cutouts, the code from this notebook would be best put into a Python script and run from the command line, saving the images or making a figure showing all of the candidate objects at once.

Discovering Hydra II dwarf

jupyter SMASH_Hydrall_interactive Last Checkpoint: Last Monday at 11:37 PM (autosaved)

Control Panel Logout

File Edit View Insert Cell Kernel Widgets Help Python 2

Code CellToolbar

Query the SMASH DR1 database

We will query the averaged photometry table from the SMASH catalog and select Field 169, which we know contains the Hydra II dwarf.

In [2]:
```
field = 169                    # SMASH Field Number to query
depth = 1                      # minimum depth
raname = 'ra'
decname = 'dec'
mags = 'gmag,rmag'
dbase='smash_dr1.object'
fid = 'fieldid'

# Create the query string.
query = ('select '+raname+','+decname+','+mags+',depthflag from '+dbase+ \
         ' where ('+fid+' = \'%d\' AND' \
         ' (depthflag > %d) and ' + \
         ' (abs(sharp) < 0.5) and ' + \
         ' (gmag is not null) and ' + \
         ' (gmag between 9 and 25) and ' + \
         ' ((gmag-rmag) between -1.5 and 3.0))') % \
          (field, depth)

print "Your query is:", query
```

```
Your query is: select ra,dec,gmag,rmag,depthflag from smash_dr1.object where (fieldid = '169' AND (depthflag > 1) and
  (abs(sharp) < 0.5) and  (gmag is not null) and  (gmag between 9 and 25) and  ((gmag-rmag) between -1.5 and 3.0))
```

We issue the query through the Query Manager, which connects directly to the database.

In [3]:
```
%%time
print "Making query"
# Call the Query Manager Service
response = queryClient.query(token, adql = query, fmt = 'csv')
df = pd.read_csv(StringIO(response))

print len(df), "objects found."
```

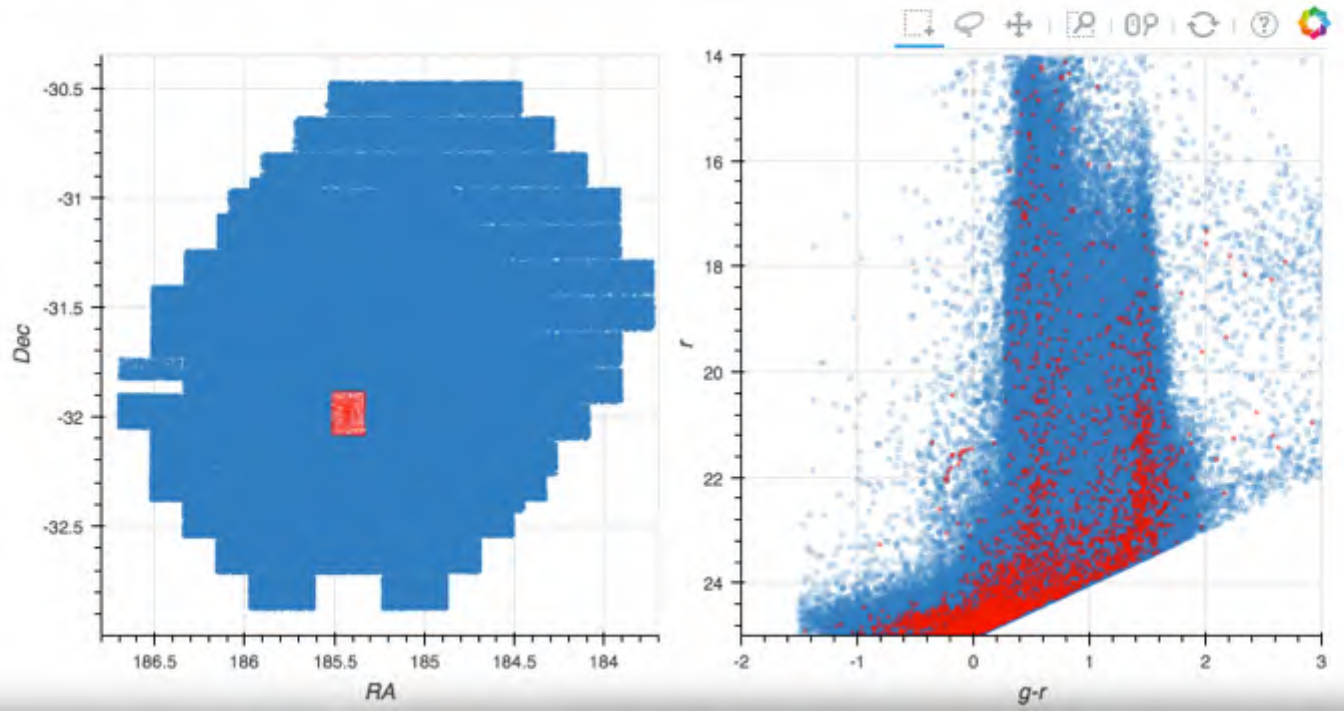
```
Making query
297788 objects found.
CPU times: user 307 ms, sys: 57.1 ms, total: 364 ms
Wall time: 12.5 s
```

Discovering Hydra II dwarf

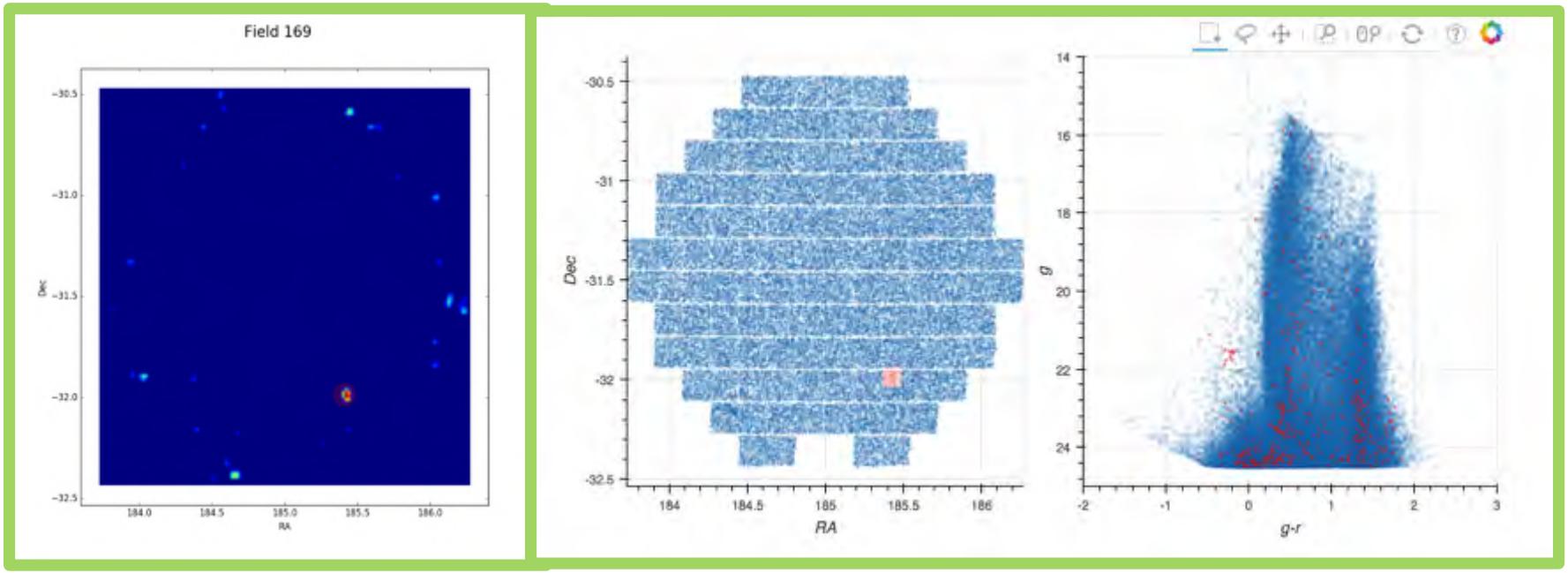
The plots

Finally, we render the plot. The figures are interactive, with ability to pan, zoom, and select samples of data that are then updated in the other plot. With the large number of points used here, the interaction can be a little slow, depending on browser and hardware. Try Box Select on the clump of points at lower left, where Hydra II is lurking.

```
In [12]: show(p)
```



Automation of workflow



From Poster 154.25

Coming in 2017

- Authentication
- Asynchronous queries and myDB through Query Manager
- Virtual storage and disk allocation
- Compute service
- Feedback? Visit datalab.noao.edu or contact us at datalab@noao.edu