Breakout summaries

Time Domain/MMA I

Moderator: Alex Razim, scribe: Azalee Bostroem

Q1: What kind of TDA anomaly (rare objects, known unknowns, unknown unknowns) we are unlikely to discover with the current LSST properties and observational strategy? What can we do to make these detections? (by Ryan Lau; follow up of the "Discovering the Unknown" session of Rubin 2023 PCW)

Key notes:

- LSST cadence discussion (survey strategy, nightly observations plan will be public -> possibility to coordinate with other surveys)
- LSST's not great for rapidly changing stochastic objects (but data volume and quality will open new possibilities) and eruptive IR variables (but Roman is good for these)
- Discussion of what LSST will be good for (e.g. red supergiants for SN outbursts)
- We need to keep the existing facilities online - LSST is not a substitute for everything else (as some TACs can think)

Time Domain/MMA I

Moderator: Alex Razim, scribe: Azalee Bostroem

Q2: What kind of ‘supplementary data’, in addition to optical LCs, promises the best cost/reward ratio for discovering new types of objects? Obtaining spectra is highly informative, but the cost of getting spectra multiple times is prohibitive. What about low-resolution spectra? Narrow-band photometry? IR, UV observations? (by Alex Razim)

Key notes:

- Following the talk by V. Placco (MW/SP I): narrow-band photometry for selecting targets for spectroscopic surveys is cool
- Low-res spectra survey for all objects up to certain brightness (e.g. mag 21; 30-40% Keck observational time?)
- IFUs, image slicers?
- Higher cadence LCs
- Target selection strategies:
    - How do we collectively coordinate what to observe?
    - Can we use unused fibers of other surveys for transients?
    - We need to keep things balanced (e.g. not to dedicate all observational time to SNe Ia)

Time Domain/MMA I

Moderator: Alex Razim, scribe: Azalee Bostroem

Q3: What kind of software or data infrastructure we are still lacking that would help us to fully utilize the synergy between already existing and upcoming surveys? (by Alex Razim)

Key notes:

- Spectroscopic surveys are not coordinated. We need a database/service to inform the community which targets are being observed/calibrated/processed to reduce duplication. Single queue for multiple facilities would be good
- Centralized database for archival data is needed (with Supernovae catalogue shutting down)
- Automatic reduction pipeline for all spectroscopic facilities
- How to deal with credit issues for data access?
- How to filter multiple alert streams?

Time Domain/MMA I

Moderator: Alex Razim, scribe: Azalee Bostroem

Q4: What kind of actions can we, as individuals, take to improve software, infrastructure, and sociology?

Key notes:

- When you're on TACs - fund software, pipelines, data reduction pipelines
- Initiate a discussion within collaborations on dealing with crediting when making data public in real time
- Scimma: observatory status coordination site (but people need to accept it -> UX, documentation, credit, publications)
- Infrastructure is of foremost importance
  - User experience
  - Documentation
  - Credit
  - Publication
- How do we reduce redundancy?
  - Multiple brokers and science platforms
  - Need to focus on connecting communities

Time Domain/MMA II

Moderator: Adam Miller, scribe: Sid Chaini

Q1: What do you think will be the biggest challenge in time-domain astronomy 5 years down the road, when we’re in the middle of the Rubin LSST operation period? (Assuming Rubin’s ToO will have found BNS via ToO observations based on LVK)

Key notes:

- Deciding which alerts are good, when triggering follow-up observations with or without a human in the loop
  - Speed and accuracy is vital
- Coordination between different collaborations and surveys
  - Even just joining data coming from different sources is difficult
  - Optimize ToOs to give back unused time
  - Move toward queue observing (especially at big telescopes)
- Amateur astronomers contribute a lot
  - Our datasets need to be amateur-friendly
  - Crediting should be taken into account
  - Standardization of how to credit contributions would be nice (ping IAU?)

Time Domain/MMA III

Moderator: Viraja Khatu, scribe: Alex Razim

Q1: One of the biggest challenges identified in TDA 5 years down the road was coordinating effectively between different communities to strike synergies between different facilities. These communities are not limited to observatories but also include academic institutions (professors, students, postdocs) as well as amateur astronomers. How can these different communities cross-apply their resources (knowledge, funding, workforce, etc.) in an attempt to solve the above challenge? What recommendations would you make as an independent advisory body tasked with coalition activities in this direction?

Key notes:

- Coordination of academic facilities is complicated; e.g. interrupting established observation/work schedule upon a request of another body is an issue for funding agencies, PIs, etc.
- Collaboration between institutions of different scale and resources: we shouldn't take talents from smaller institutions, but bring research to them
- 'Nothing brings people together as money': a funding call that demands collaboration forces institutions to collaborate
- Can we organize brainstorming events to define common goals, so that funding agencies didn't have to choose between multiple non-compatible projects?

Time Domain/MMA III

Moderator: Viraja Khatu, scribe: Alex Razim

Q2: How can we measure the potential to discover novelties? For a lot of more conservative science (low risk medium reward) we can easily quantify survey performance on paper, but for “exploring the transient universe” and in particular for the discovery of novelties, high risk high reward science, we cannot, which means when strategic decisions are made we have a harder case to make.

Key notes:

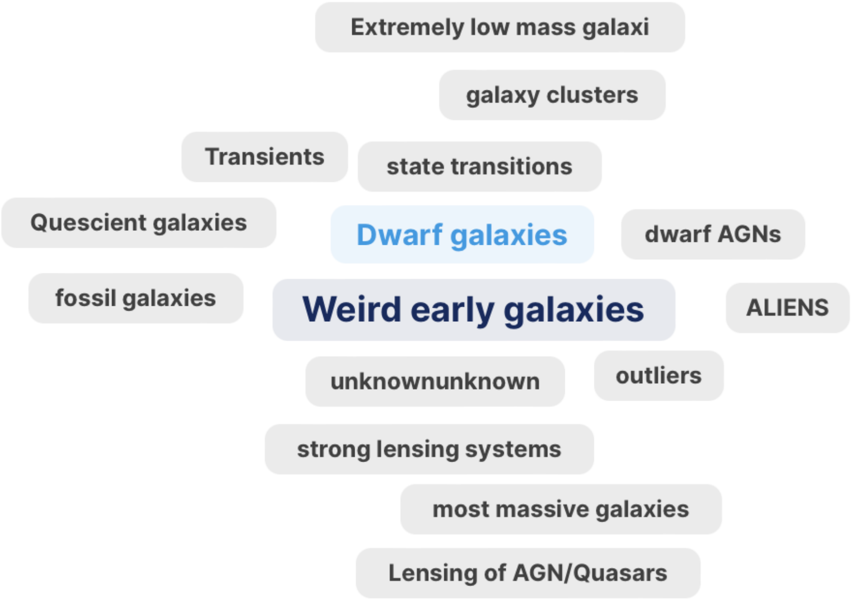
- ‘Measuring the potential’ is a dangerous thing; there is always a way to trick a KPI. Can we organize anomaly-detection-oriented workshops/hackatons to change the mindset and ‘orient’ people towards novelty discovery?
- How do we convince those who allocate money and time that high-risk high-reward science cases are worth it? A TAC-member opinion: do not assume TACs to be completely conservative
- A positive phrasing for high-risk proposals: “we explore the parts of the parameter space which we didn’t probe before to see if there is something new”
- Suggestion for TACs: during calls for proposals openly encourage high-risk proposals
    - Dedicated high risk time set aside?

Galaxies/Cosmology

Chair: John Wu
Scribe: Stephanie Juneau & Antonella Palmese

Galaxies/Cosmology I

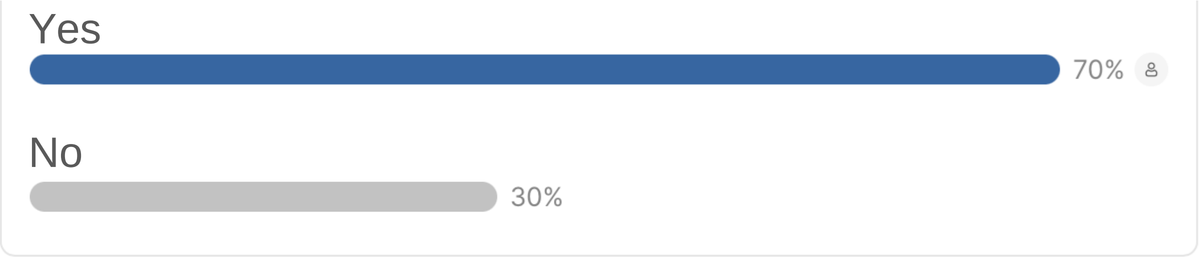
What are the rare gems that will revolutionize our understanding of galaxies and cosmology?



- Weird early galaxies that may break LCDM
- Dwarf galaxies to learn about galaxy formation, DM, and reionization.
- Short-lived transient states in galaxy evolution.
- Unknown unknowns.

Galaxies/Cosmology I

Will we make any “big” discoveries in galaxies and cosmology in the next decade?



No:

- Most progress is incremental.
- Large discoveries more likely in cosmology than galaxies.

Yes:

- Dark Matter still total unknown.
- Hubble tension, DESI evolving Dark Energy?
- New technologies usually lead to new discoveries.

Galaxies/Cosmology I

What is the best role for ML/AI to enable breakthrough discoveries?

- Remove the boring stuff
- Optimize pipelines
- Finding more examples of unique objects
- Find gaps and transitions
- Cross-matching
- Hypothesis generation?
- Dimensionality reduction

Galaxies/Cosmology I

What is the best role for ML/AI to enable breakthrough discoveries?

- Remove the boring stuff
- Optimize pipelines
- Finding more examples of unique objects
- Find gaps and transitions
- Cross-matching
- Hypothesis generation?
- Dimensionality reduction

ML/AI is a tool. Discoveries come from people.

Chair: Burçin Mutlu-Pakdil
Scribe: Jeff Carlin

Galaxies/Cosmology II

Are you ready (feel prepared) for the LSST era?

No:

- Don't feel prepared to ask science questions.
- Unsure about navigating data products and tools/skills.
- Increased workload.
- Individual vs. community projects: is there a change in perspective or working style?

Yes:

- We have a lot of the necessary tools ("just" need to scale up).
- Community is (slowly?) getting used to going to a science platform or portal
- Already done in other realms (e.g., MeerKAT, ALMA) where they've already got "big data".
- "We're just too curious as a field to just give up and say we're not ready to deal with that volume of data."

Galaxies/Cosmology II

What is "the problem" that we should focus on as a community?

- Galaxy ecosystems, linking black holes, galaxies, and the cosmic web.
- Find and characterize a "complete" sample of local dwarf galaxies.
- Dark energy, dark matter, Hubble tension.

What are the biggest challenges we will face in the LSST era?

- Survey uniformity: Galactic cirrus, scattered light, background subtraction, source shredding, etc.
- Need to characterize the data set via synthetic source injection.
- Learning curve, adjustment to new ways of doing science
- Avoiding an escalation of competition and effort commensurate with the escalation in data volume.

Galaxies/Cosmology II

Are “rare gems” actually that important?

- You're not going to convince someone you have upended a major theory with a single object.
- Many of the most-cited early SDSS papers are actually studies that used large samples of objects rather than “rare gems.”
- But: finding the one interesting object motivates finding more similar objects (and funding searches for more objects).

Galaxies/Cosmology II

Are “rare gems” actually that important?

- You're not going to convince someone you have upended a major theory with a single object.
- Many of the most-cited early SDSS papers are actually studies that used large samples of objects rather than “rare gems.”
- But: finding the one interesting object motivates finding more similar objects (and funding searches for more objects).

“More money has been made from mining coal than from mining diamonds.”*

*unsubstantiated claim

Chair: Mike Jones
Scribe: Stephanie Juneau & Antonella Palmese

Galaxies/Cosmology III

What tools do we need to develop to address the challenges of the LSST/Big Data era?

- We have a tendency to use the tools we know, rather than identifying/inventing the right tool for the task.
- Simulations to make robust comparisons to the enormous and varied galaxy sample that LSST/DESI/Euclid/Roman will deliver.
- The ability to go back to the raw data a reprocess if needed.

Chair: Mike Jones
Scribe: Stephanie Juneau & Antonella Palmese

Galaxies/Cosmology III

What tools do we need to develop to address the challenges of the LSST/Big Data era?

- We have a tendency to use the tools we know, rather than identifying/inventing the right tool for the task.
- Simulations to make robust comparisons to the enormous and varied galaxy sample that LSST/DESI/Euclid/Roman will deliver.
- The ability to go back to the raw data a reprocess if needed.

The dream: A completely modular workflow where you can drag-and-drop the tool you want for every step.

Chair: Mike Jones
Scribe: Stephanie Juneau & Antonella Palmese

Galaxies/Cosmology III

What tools do we need to develop to address the challenges of the LSST/Big Data era?

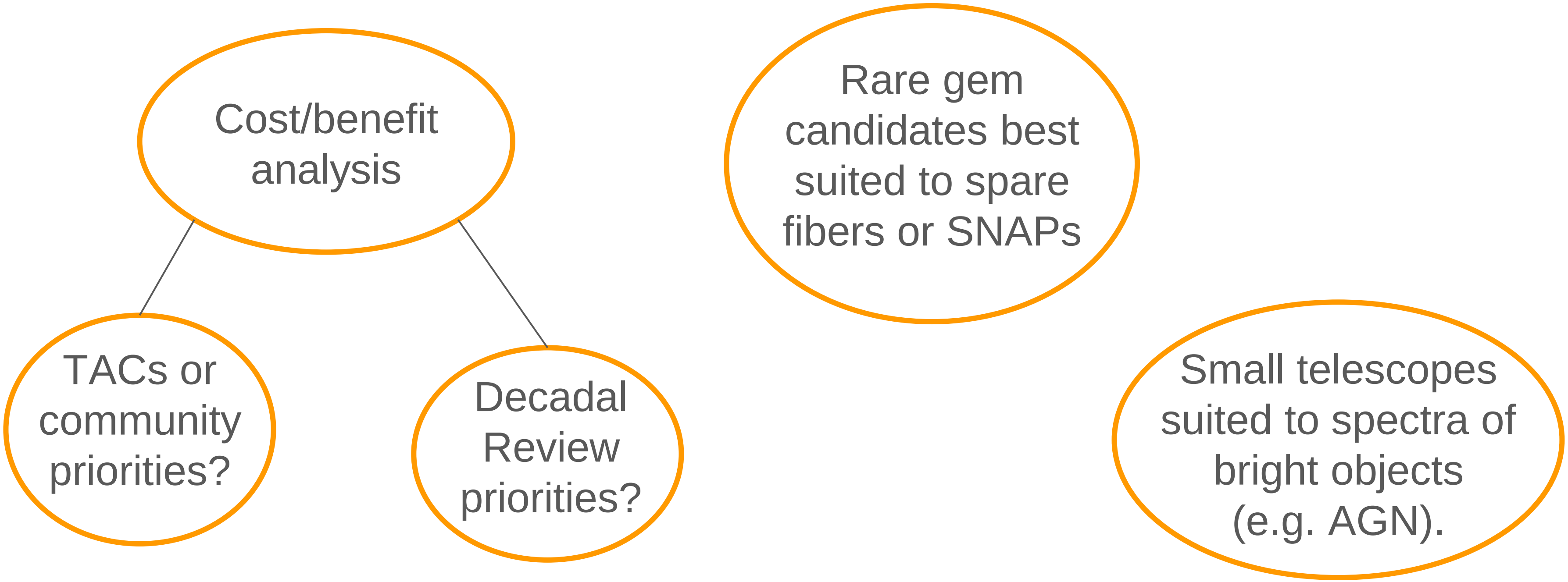
- We have a tendency to use the tools we know, rather than identifying/inventing the right tool for the task.
- Simulations to make robust comparisons to the enormous and varied galaxy sample that LSST/DESI/Euclid/Roman will deliver.
- The ability to go back to the raw data a reprocess if needed.

The dream: A completely modular workflow where you can drag-and-drop the tool you want for every step.

The dream II: Gpc^3 fully hydro simulation suite with varied baryonic physics.

Galaxies/Cosmology III

How are we going to decide what to (and what not to) follow up?



Galaxies/Cosmology III

What big data sets will we want to combine to maximize discovery potential and where/how will this happen?

- We all agreed on the “obvious” choices e.g. **LSST, DESI, Roman, Euclid, eROSITA, CMB-S4**.
- At the catalog level much of this will likely be possible on the Rubin Science Platform or equivalents for other projects.
- Different platforms will likely have different protocols: Maybe we can learn from services like **CDS** and **IVOA**?
- For imaging there will likely be logistical issues with where different data sets are located, complicating bringing compute to data.
- However, single largest barrier is likely to be **data access rights**.
- May need to wait for public releases and hope these have all that’s required.

Galaxies/Cosmology Summary

- We are excited for LSST and the Big Data era! However, there are some challenges we still need to work out how to resolve.

- Rare gems a vital to driving curiosity and motivating follow-up investigations.

- However, the “coal” is very important too.

- We expect most progress in galaxy evolution to be incremental, but cosmology may be a different story.

- ML/AI will be an essential tool, but it is still a tool, the discoveries will come from people.

Milky Way/Stellar Populations I

Chair: Eric Bell
Scribe: Sergey Koposov

Given the rapid growth in multi-dimensional astronomical data, how can machine learning be harnessed to enhance the process of discovery, including distinguishing interesting anomalies and finding patterns, reducing false positives, managing domain knowledge for rare object classification, using those to build physical intuition?

- Use of **latent variable**
- **Emulate simulations** based on limited high-res data
- Incorporate domain knowledge into machine learning with **supervised approach**. Can work towards connecting unsupervised and supervised approaches
- Rigorous testing against noise with tailored ML algorithms. Difference between image vs catalog detections. **The closer to the data that you are the better you understand the errors.**
- **Educational needs** : integrating machine learning and statistical training in astronomy

Milky Way/Stellar Populations II

Chair: Denija Crnojevic
Scribe: Amandine Doliva-Dolinsky

Tutorial session on the use of the NOIRLab Source Catalog through Astro Data Lab

- Jupyter notebook is in the breakout slack channel !

Rare Gems NOIRLab Source Catalog (NSC) Tutorial

```
[1]: __author__ = 'David Nidever <dnidever@montana.edu>, Pol Massana <pol.massana@montana.edu>'
     __version__ = '20240521' # yyyymmdd
```

1) Installing packages and imports

Pip installing `Leavitt` should get you everything you need.

```
pip install leavitt
```

```
[2]: # Some imports
     import numpy as np
     import matplotlib.pyplot as plt
     from matplotlib.colors import LogNorm as LN
     %matplotlib inline
```

2) NSC Introduction

The NSC is a crowd-sourced survey of the sky using all the public exposures in the NOIRLab Astro Data Archive.
It covers 85% of the sky, contained 3.9 billion objects and 68 billions measurements.
1.5 billion objects have 10+ measurements.

More information:
https://datalab.noirlab.edu/nscdr2/index.php
https://ui.adsabs.harvard.edu/abs/2021AJ....161..192N/abstract

Milky Way/Stellar Populations II

Discussion around the NOIRLab Source Catalog

- Some aspects **transferable** to the upcoming LSST (Legacy Survey of Space and Time) project
- The database is underutilized
- Excitement about discovering variable stars and utilizing machine learning to tackle complex science questions with this catalog.
- Challenges with **star/galaxy separation** at faint magnitudes, the importance of **quality flags** for identifying outliers.
- Integration of this data with future datasets like those from LSST (funding and collaboration issues?)
- **Preparatory work** : using templates
- End note : **a call for a Noirlab catalog hackday over zoom**

Milky Way/Stellar Populations III

Chair: Amandine Doliva-Dolinsky
Scribe: Jeff Carlin

In this Big Data era, we will need a lot of follow-up of discoveries. What instrumentation/facilities will be needed, and how do we prioritize the efforts?

- S**ample size depend on the question** : study of a population or search/analysis of rare gems. With uniform, well-characterized survey + oversubscription of follow up: **get away from understanding each object well to thinking about the population?**
- **Lack of spectroscopic power for follow up** (a lot of cases requires 10-meter class telescope) -> there might be other option (narrow band photometry)
- Follow-Up Timing: Low-mass galaxies are not transient, allowing flexibility in follow-up timing.
- **Equity in Access:** people at institutions with proprietary access to the large aperture facilities that get to follow things up -> **Push for collaboration**.

Milky Way/Stellar Populations III

What have you been excited about at this conference? What are the highlights, what are things that can be solved with LSST (and other big data surveys)? What are the challenges that you see?

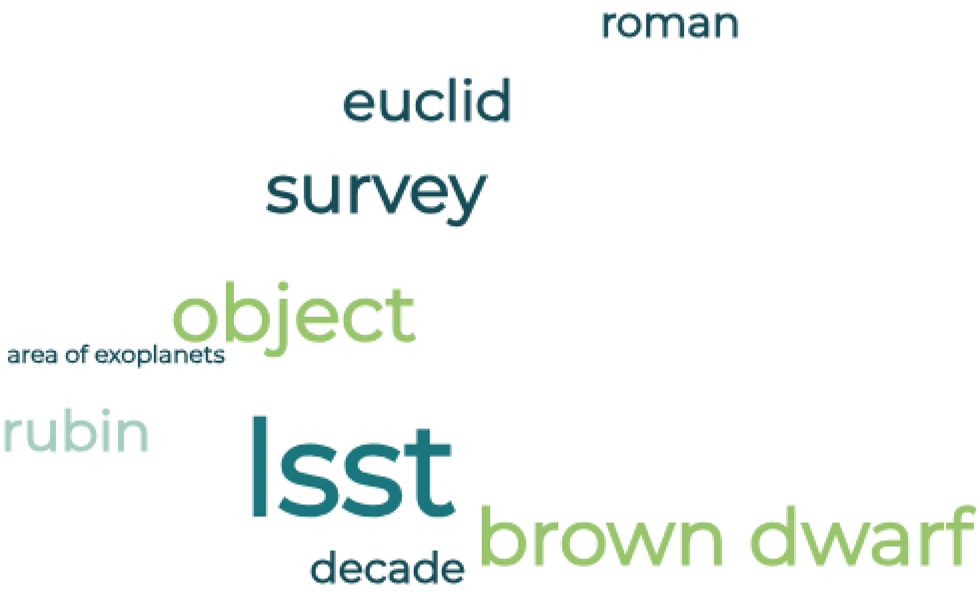
- **Use of ML as a tool** in this new era with awesome possibility ! A few challenges : learning for students but also researcher, computing resources. **We have to bring along the different communities** (even reach out to industry): this conference was a great example on how to do that !

- **Combining LSST with other datasets** for analyses: Fundings and politics BUT Rubin+Euclid has produced lots of fruitful conversations and Rubin+Roman conversations are ongoing.

- How do we enable analyses with multiple joined datasets on the science platforms that we provide to the community? This may be another case where if additional compute/storage resources are needed

Remember – it’s an amazing time to be an astronomer!

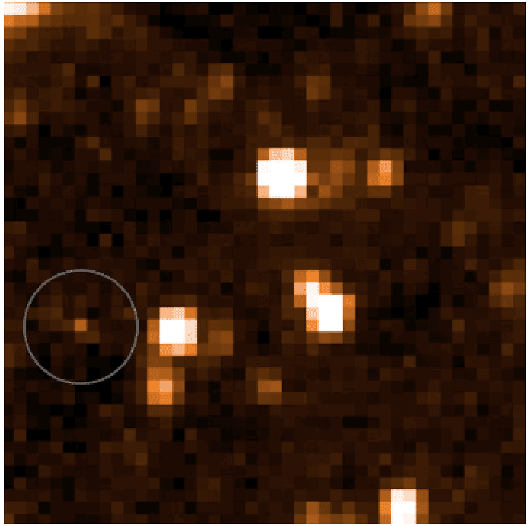
exoplanets / moving objects



scope: exoplanets, solar system, brown dwarfs, proper motions

exoplanets / moving objects

- Day 1: discussed the landscape of current/upcoming surveys/data of relevance to this science theme
- Day 2: discussion of tools including brokers, archives, science platforms, machine learning, and citizen science
- Day 3: forward-looking discussion about the most exciting discoveries/advancements to come for this science theme

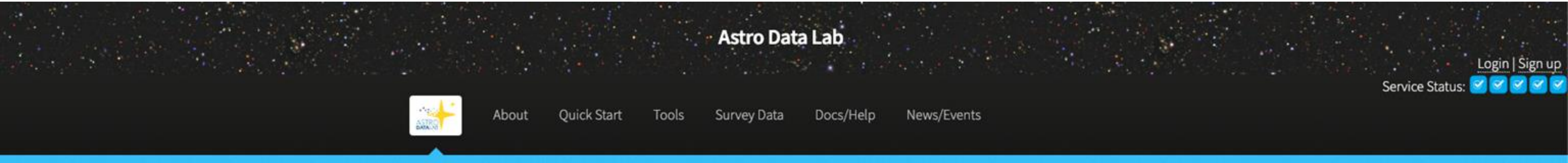


WISE 1534-1043 a.k.a. The Accident (discovered by citizen scientist and Rare Gems attendee Dan Caselden)

moderators: Sarah Caswell (Monday), David Trilling (Tuesday), Philip Choi (Wednesday)

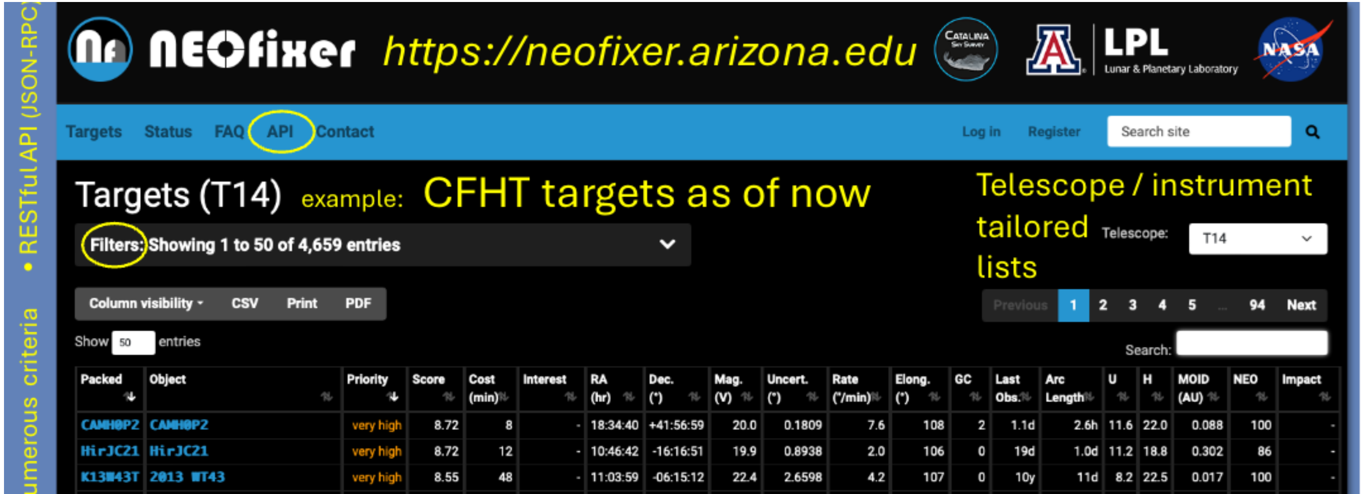
exoplanets / moving objects: takeaways (1/4)

- our science area requires the ability to effectively work across different surveys, ideally without every researcher reinventing e.g., cross-match tables on their own
- science platforms such as NOIRLab Astro Data Lab and Rubin Science Platform seem highly relevant here
  - will there ever be one central place to analyze all the big upcoming astro data sets in an efficient, colocated manner?
- cross-matching for moving objects can be less straightforward than for stationary objects like galaxies



exoplanets / moving objects: takeaways (2/4)

- lots of interest in brokers for moving objects (both solar system and Milky Way)
- would definitely be nice to the extent that brokers can include features geared toward moving objects, in addition to static flux variable sources
- some existing examples of such functionality were raised:
  - ‘fink’ broker includes solar system object support
  - neofixer (see below)



exoplanets / moving objects: takeaways (3/4)

- strong interest in both machine learning and citizen science, plus the combination of these two
- Daily Minor Planet is an innovative example posting new data to Zooniverse (roughly) every night
- Backyard Worlds: Cool Neighbors is an example currently combining machine learning with citizen science
- ML+CitSci: interest in further ‘closing the loop’ by using volunteer classifications as training data
- not clear whether citizen science will/won’t naturally scale to accommodate the size of upcoming data sets e.g., Rubin/LSST

exoplanets / moving objects: takeaways (4/4)

- brainstormed the most exciting ‘rare gems’ on the horizon within this science theme
    - trans-Neptunian objects
    - centaurs
    - Earth trojans
    - interstellar objects passing through the solar system
    - tiny moons of solar system planets
    - O(10x) increase in solar system minor body census
    - Earth-like exoplanet(s) in the habitable zone (obligatory)
    - Jupiter-temperature (or colder) brown dwarf(s)
    - new brown dwarfs among nearest systems (closer than Proxima?)
    - hypervelocity stars