



# Science platforms in the age of time-domain astronomy

---

Robert Nikutta, Astro Data Lab / NOIRLab  
*robert.nikutta@noirlab.edu*

*November 11, 2023 / IVOA Interop Tucson*

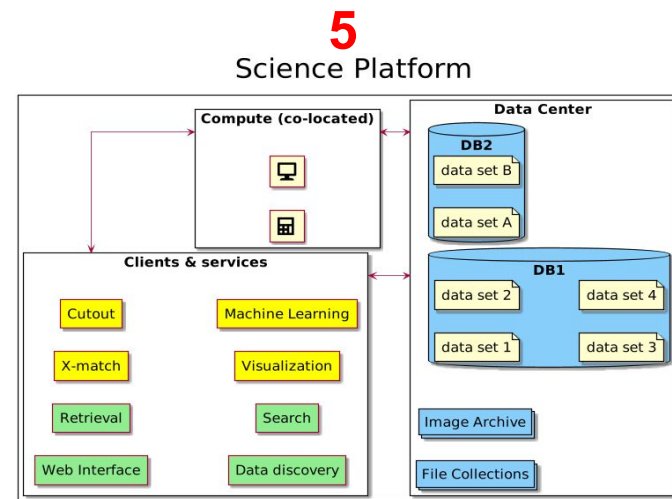
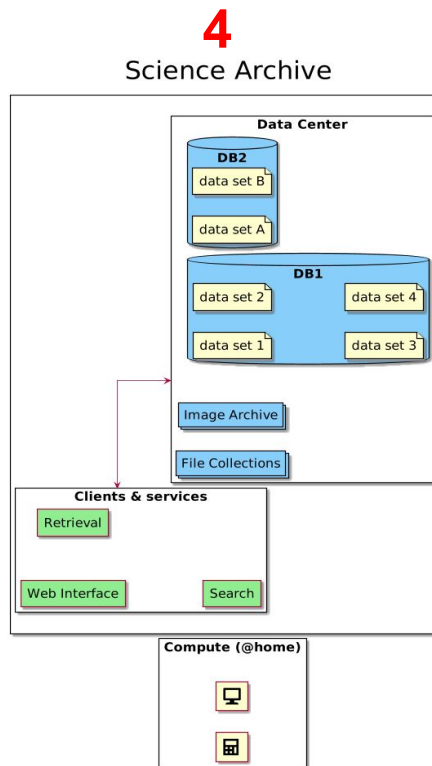
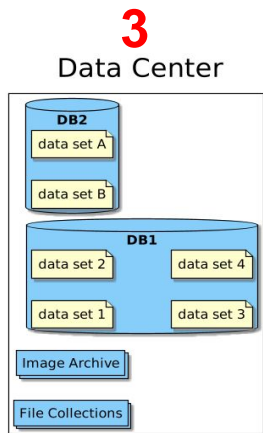
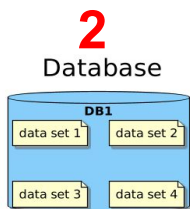
# Caveats about this presentation

- I'm not much of a TD practitioner!
- I don't write observational proposals
- I don't have watchlists on your favorite alert brokers
- I'm not even a member of this TDIG (but I should join!)
- I've pulled some figures and slides from my smarter colleagues
- Any opinions are my own (and I shall own them alone)

But:

- I come from the perspective of SPs (Astro Data Lab @NOIRLab)
- They have diverse community of users (science interests, professional seniority, culture, age, gender, nationality, available resources at their home institutions, etc.)
- All want to do science, including TD astronomy
  - *And often this means: whichever way they can accomplish their goals*
- Majority of practicing astros don't "care" about standards (as in "they don't think about them much")
  - *In a strong sense: standards are for data service providers. Users exercise them through abstract clients.*

# From dataset to complex platform



And that's just already observed data...

## So, a Science Platform is... (one definition)

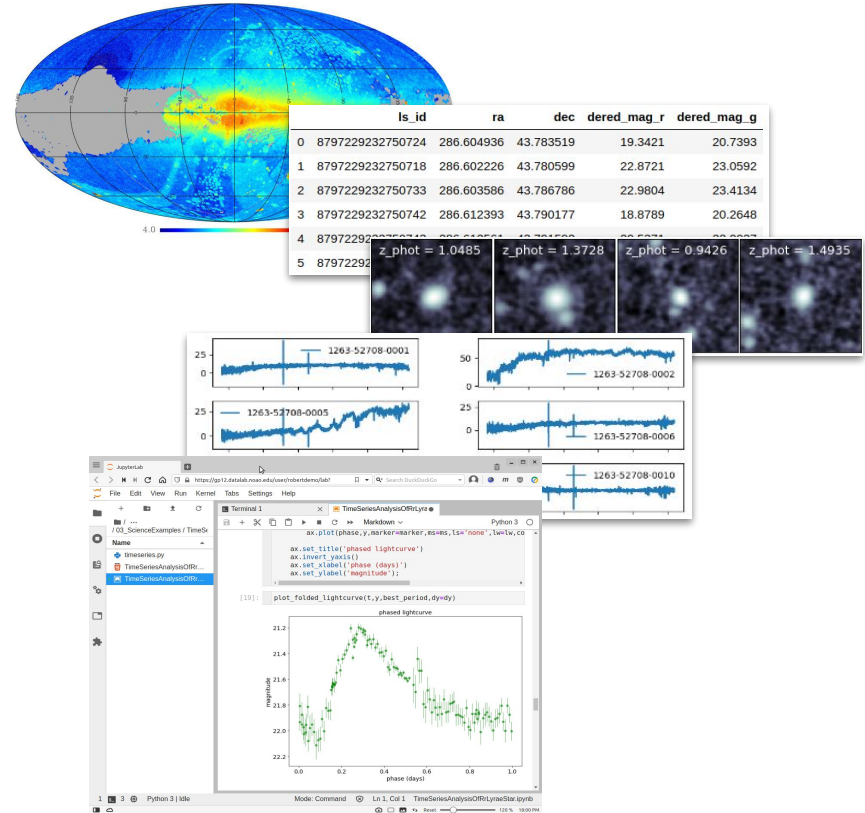
- Where big data and compute are co-located = archive + CPU/GPU
- Data services connecting them
- Front-ends and public APIs exposing them

Some common boundary conditions:

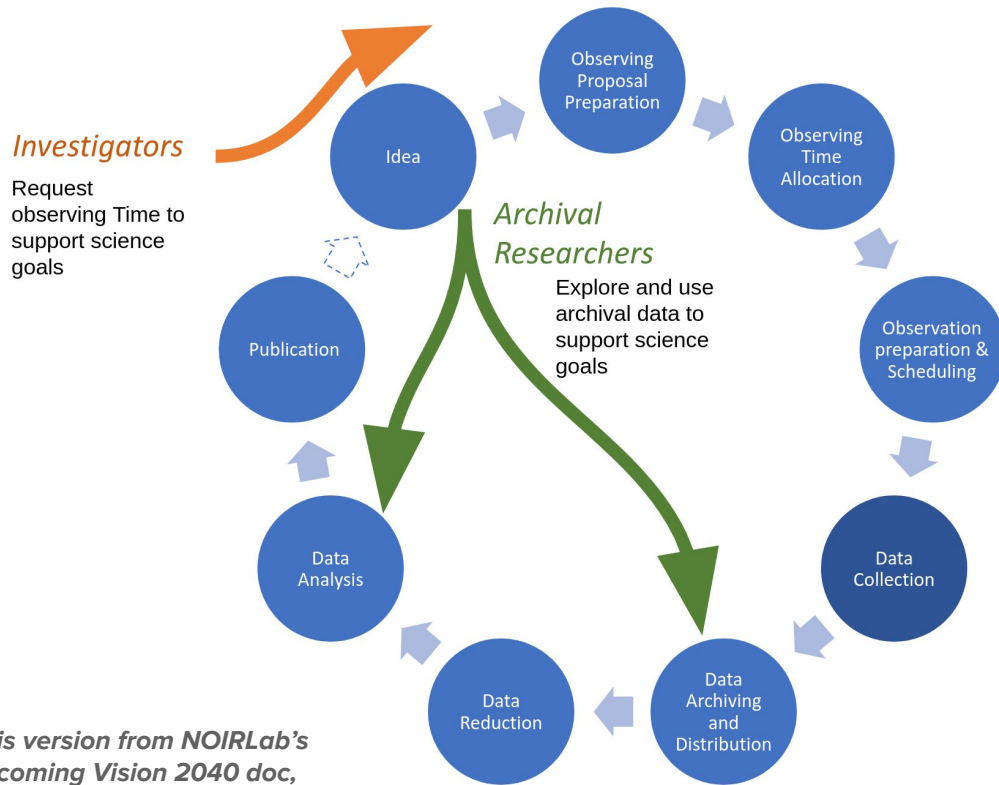
- Usually SP defined per-mission or per-organization, or per-country
- Usually own AuthN & Z (though federation does work)
- Usually (but not always) open to a specific (and limited) user community
- Usually limited or tightly controlled resources
- Accessing external compute hard (technically) or impossible (politically)

# Example: Astro Data Lab

- Catalogs (ivoa:tap, ivoa:scs, ivoa:ucd)
- MyDB / User DB tables (ivoa:tap)
- Images (ivoa:sia)
- Spectra (custom SPARCL service)
- Remote user file storage (ivoa:vospace)
- Compute: NB servers
- Time domain:
  - Measurement tables
  - ANTARES alert broker filter dev kit
  - TD example NBs
- And 2,800+ users to keep happy...



# Data lifecycle



## Investigators

Request observing Time to support science goals

## Archival Researchers

Explore and use archival data to support science goals

*This version from NOIRLab's upcoming Vision 2040 doc, by Marie Lemoine-Busserolle*

Time domain adds epicycles, e.g.:

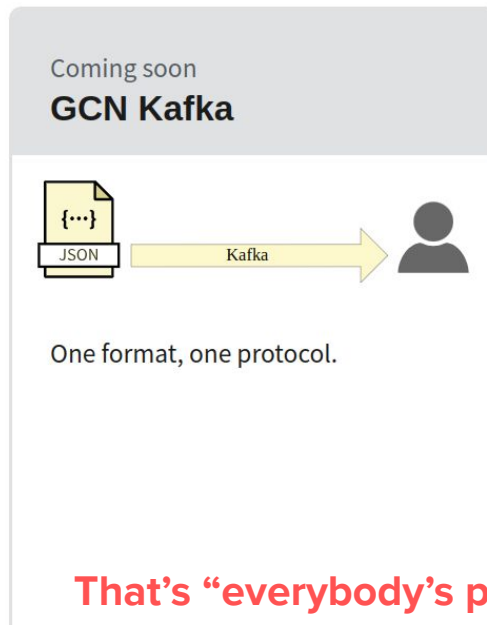
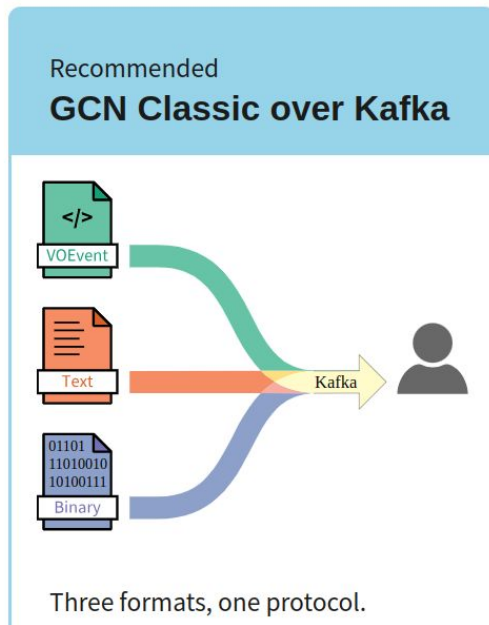
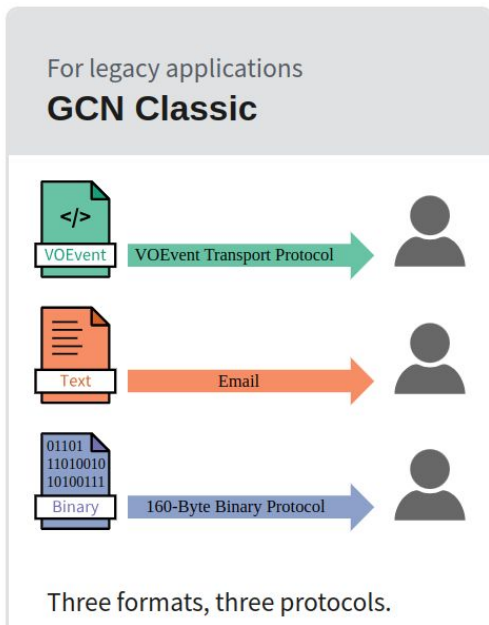
- Near-live monitoring for transients
  - *Filtering alert streams*
  - *Might be MMA*
- Classification of alert
  - *Too many approaches to list*
- Importance weighing / prioritization
  - *Needle in haystack, or YASN?*
- Must have ToO proposal on standby
  - *And likely at multiple facilities*
- Trigger appropriate follow-up obs
  - *Right facility*
  - *Right time*
  - *Right conditions*
- Loop in collaborators, other facilities

# The vision vs reality in TD astro (the bad)

- IVOA as a standardizing entity is too slow for the current pace of development
- Funding levels for many (non-datacenter level) projects are chronically low  
→ *E.g., GOATS core team = 1 scientist + 1 developer*
- Adopting “custom” standards is then a luxury  
→ *So they run with off-the-shelf stuff proven to work, e.g., Kafka, JSON, RESTful APIs*
- Tech like Kafka and JSON has a **\*much\*** larger development community than astronomy  
→ *Much easier to find help, recipes, bug reports, solutions, and actual implementations*
- Any S/W engineer entering astro is immediately stopped.  
→ *No REST, no JSON response for service calls? = No framework out there to work with.  
(See Gregory D-F remarks here and in Bologna)*
- It's a race to the top (who does some new science first)  
→ *But from the POV of standards, it's really a race to the bottom (everybody siloes themselves in)*

# Alert formats and transport protocols

From the new General Coordinates Network site: <https://gcn.nasa.gov>



IVOA is being phased out here...

...but could help here.



## The vision vs reality (the good)

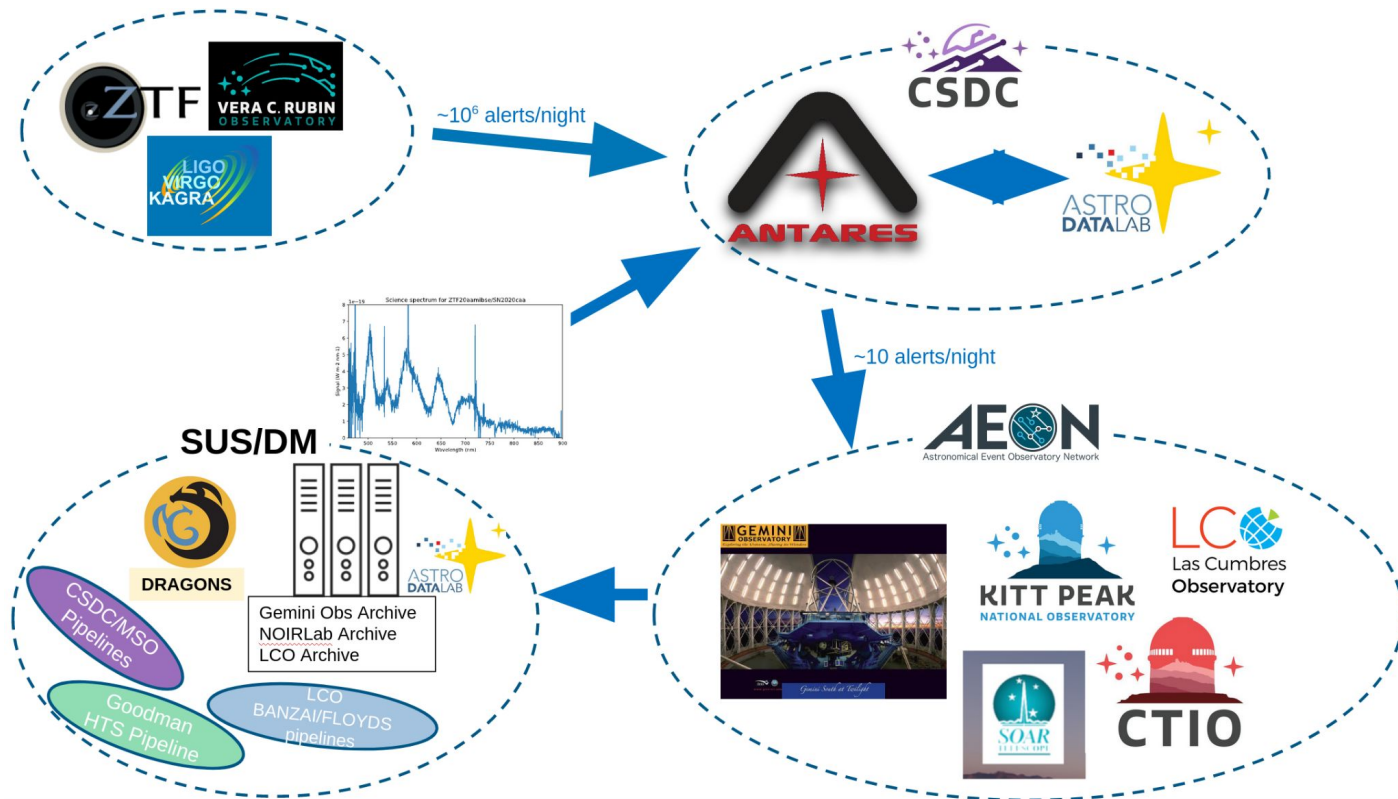
- High pace of development usually means field isn't mature yet  
→ *It might be good to let it settle a bit*
- The largest gorilla sets the “standards”  
→ *Whatever Rubin, SKA, etc., adopt*  
*(but possibly “whatever FAANG adopts”)*
- Post-facto standardization?  
→ *Once field settles, stability of protocols and interfaces will be sought-after*

**Doesn't mean “Wait and see”**

**but**

**“Get biggest players to agree on formats & protocols now” (IVOA or otherwise)**

# Time-Domain cycle (NOIRLab example)



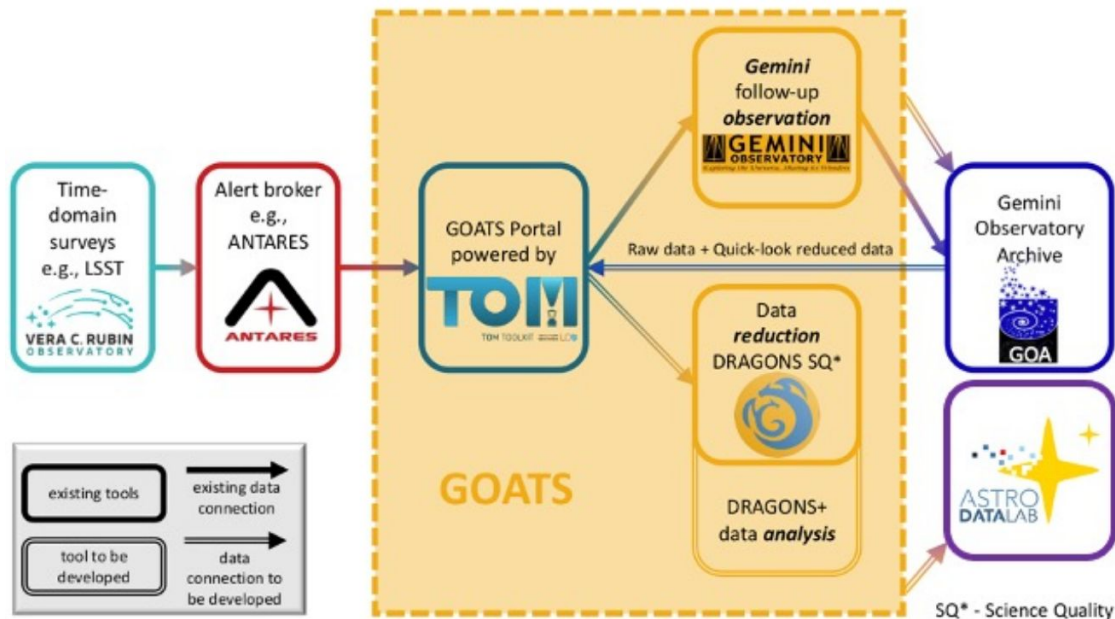
From Monika Soraisam  
(GOATS PI)

# The future at NOIRLab: GOATS



## Gemini Observation and Analysis of Target Systems

- Automatically trigger follow-up observations
- Reduce data via pipelines (auto or interactive)
- Automatically transfer all raw and reduced data to archive



From Monika Soraisam  
(GOATS PI)

# SPs as “glue” of all that

- All functionality is a “pip install” away, e.g. clients to filter alert broker streams
- Big data co-located (e.g., for MMA)
- Compute next to it (e.g., for classification work)
- User storage (files and DBs)
- Integrate TOMs, meta-TOMs, e.g., GOATS
  - *Interface with Data Lab planned for 2025/2026*
- Integrate data reduction pipelines (e.g., DRAGONS)
- Analysis frameworks and compute (e.g., Jupyter)
- Plus all of Python, Viz, collaboration, etc.

***Thank you!***

