

The ANTARES Project

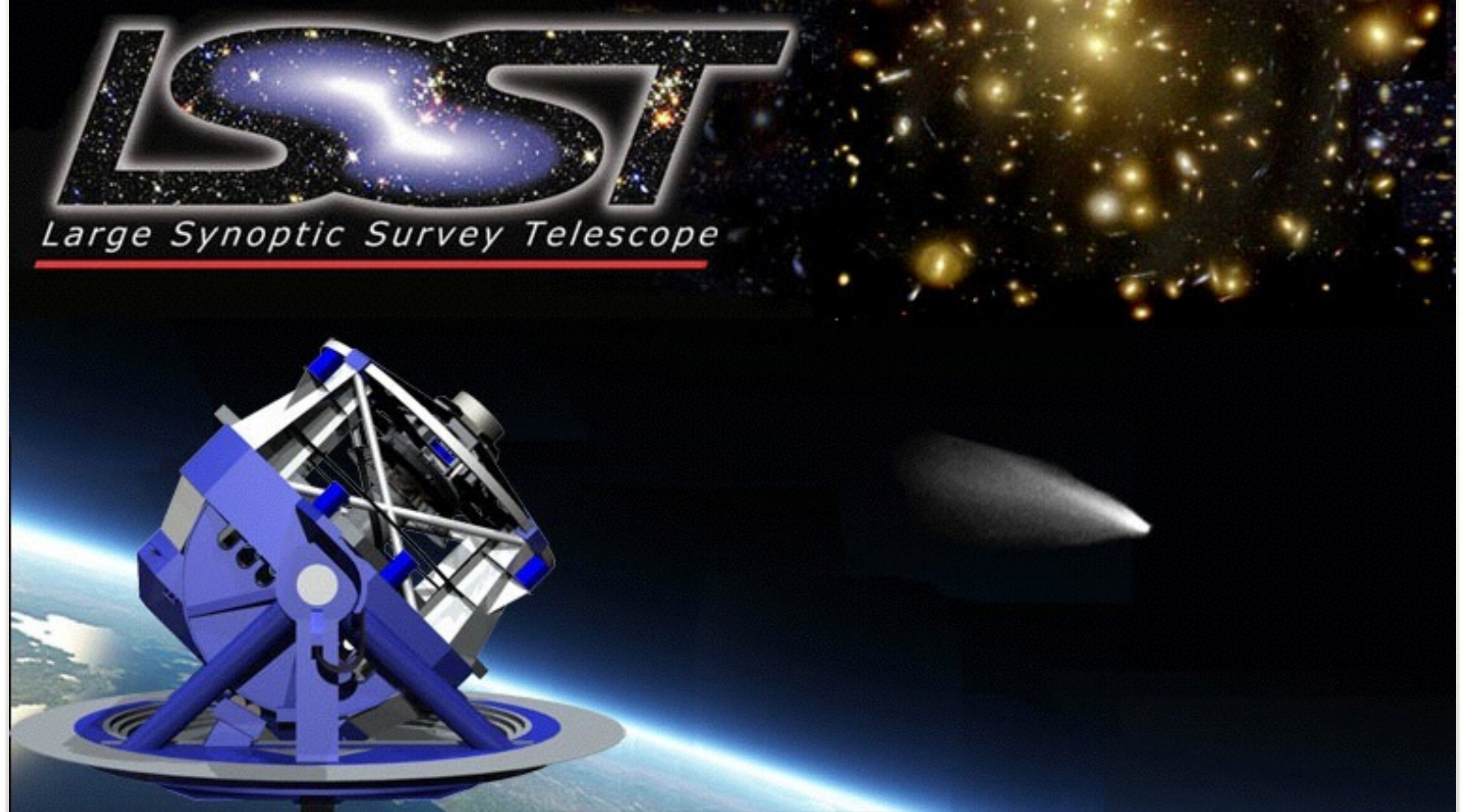
Hunting the rarest of
the rare

Gautham Narayan
NOAO/UA

Rick Snodgrass, John Kececioglu, Abi Saha, Tom Matheson, Rob Seaman, Tim Jenness

Grad Students: Jackson Toeniskotter, Shuo Yang, Zhe Wang, Jen Dempsey

NSF INSPIRE proposal (IIA-1344024)



<http://www.symmetrismagazine.org/article/march-2015/a-telescope-that-tells-you-when-to-look-up>

LSST is the elephant
in the corner

LSST understands the need
for Alert Brokers

We are building an alert
broker!

“Drink deeply from the firehose.”

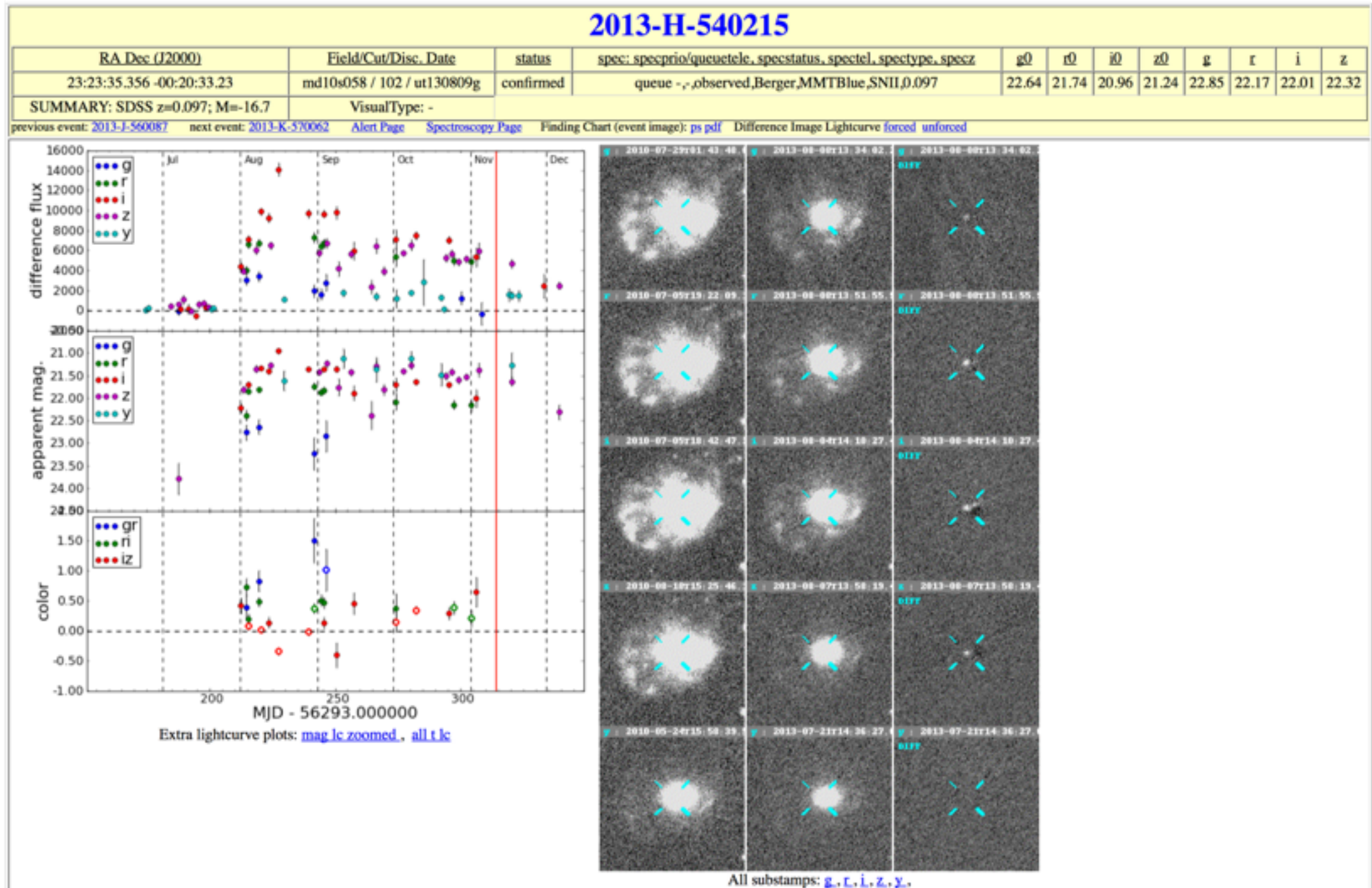
–Eugene Magnier

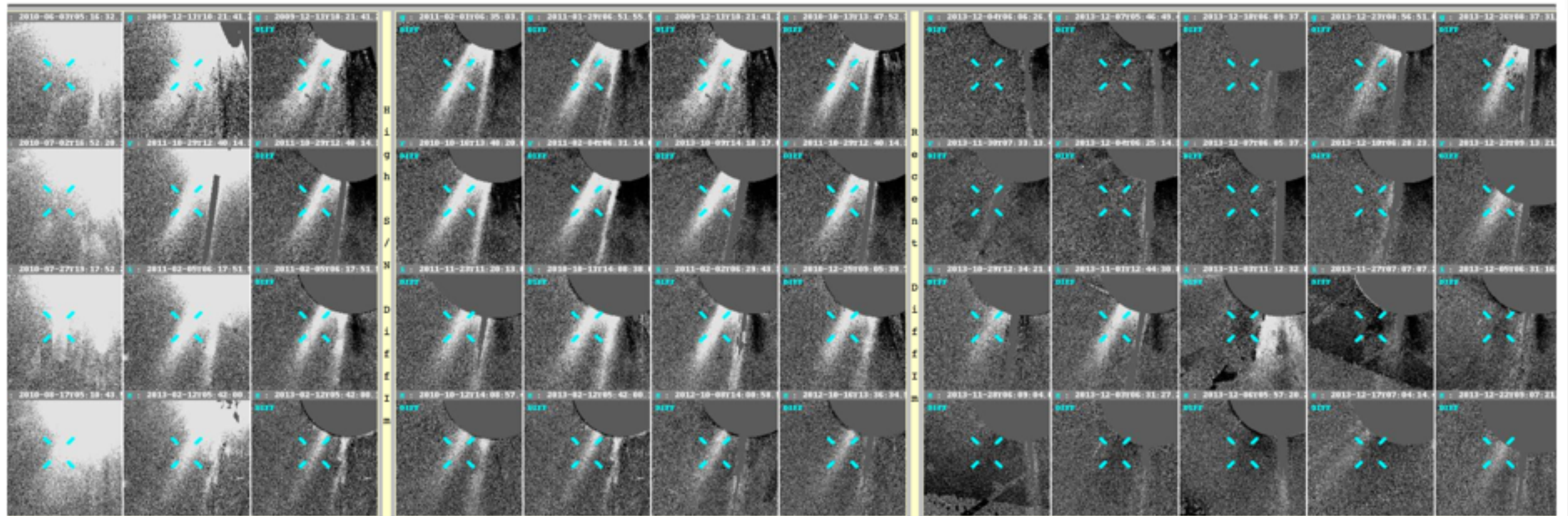
*Email to the IPP Users Group on commencement of PS1 “Demo Month”
data processing*

02/11/2010

The Petaflared

- ❖ Triangulum
- ❖ Orion
- ❖ Great Orion
- ❖ Lagoon
- ❖ of
- ❖ ~
- ❖ P1



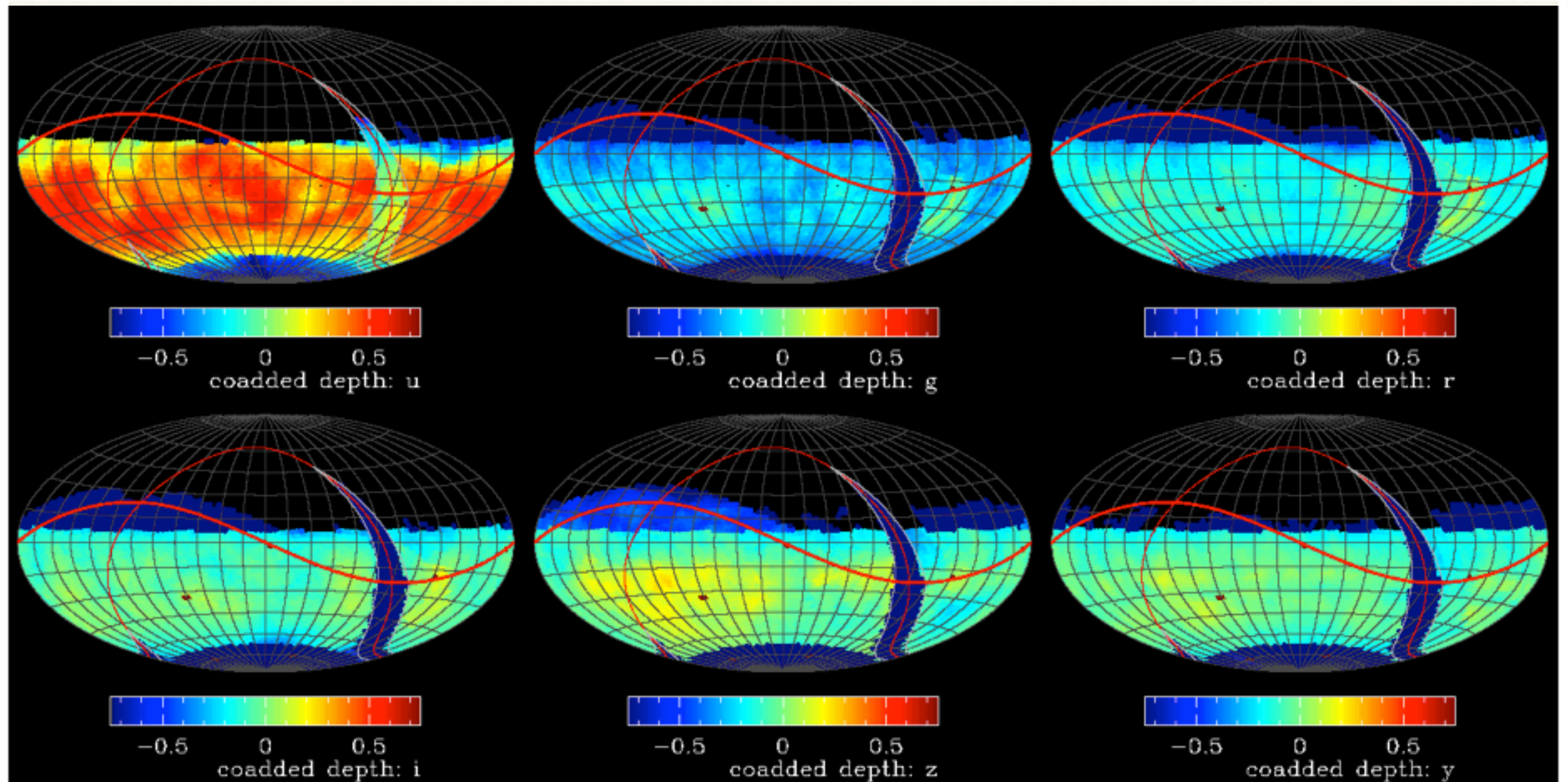


Number of Artifacts are much larger than number of Alerts

The Alert Stream is Dirty

ML being used for artifact rejection - work on using random forests. Already being used by PTF survey (Brian Bue)

See Guillermo Cabrera (U. Chile)



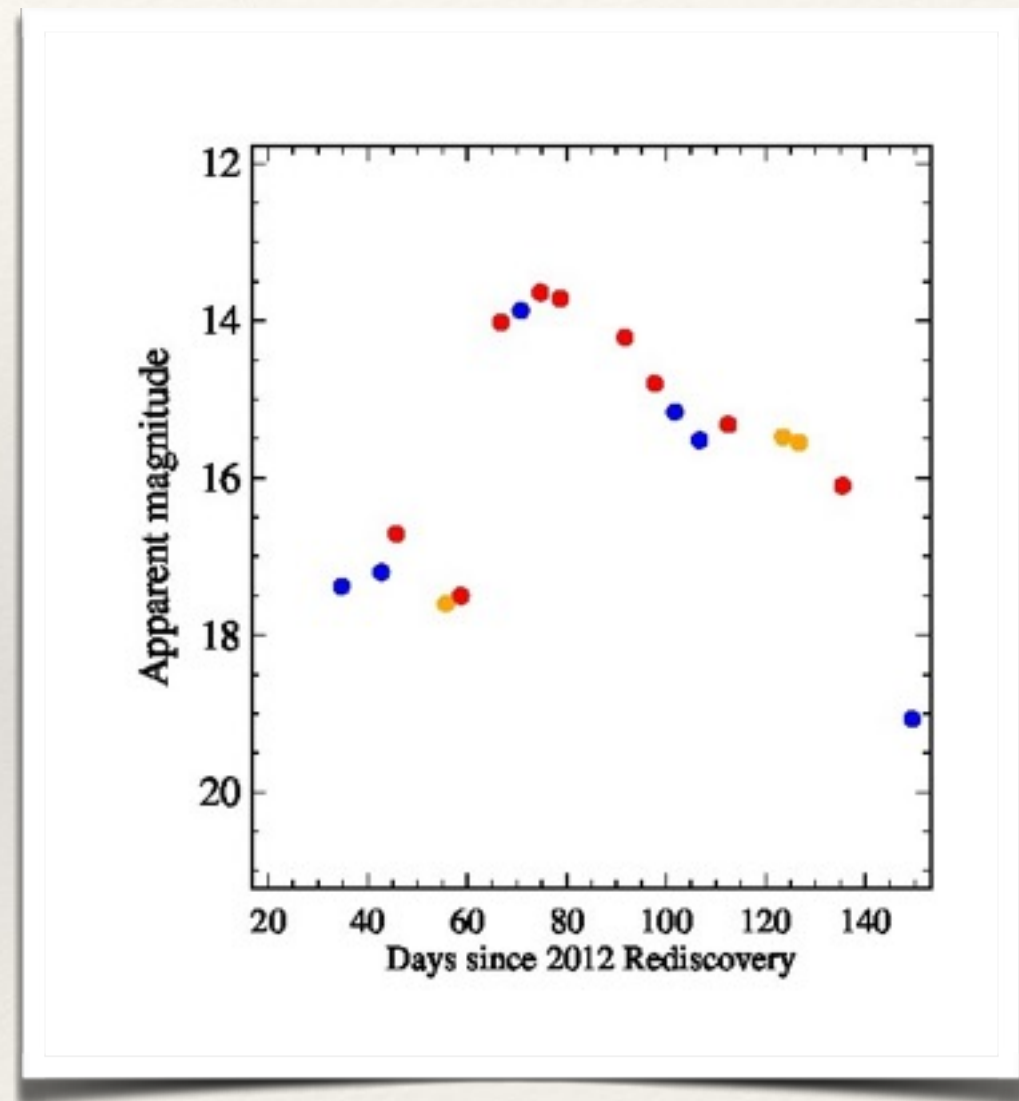
Cadence

Every year is getting shorter
Never seem to find the time
Plans that either come to naught
Or half a page of scribbled lines

-R. Waters+, 1973
The Dark Side of the Moon

Why should you care about cadence?

- ❖ LSST has only simulated a few cadences
- ❖ Metrics Analysis Framework (MAF) to assess different cadences
- ❖ **You** write metrics to optimize for the science you care about, LSST runs simulations



A big problem, with a side of challenges

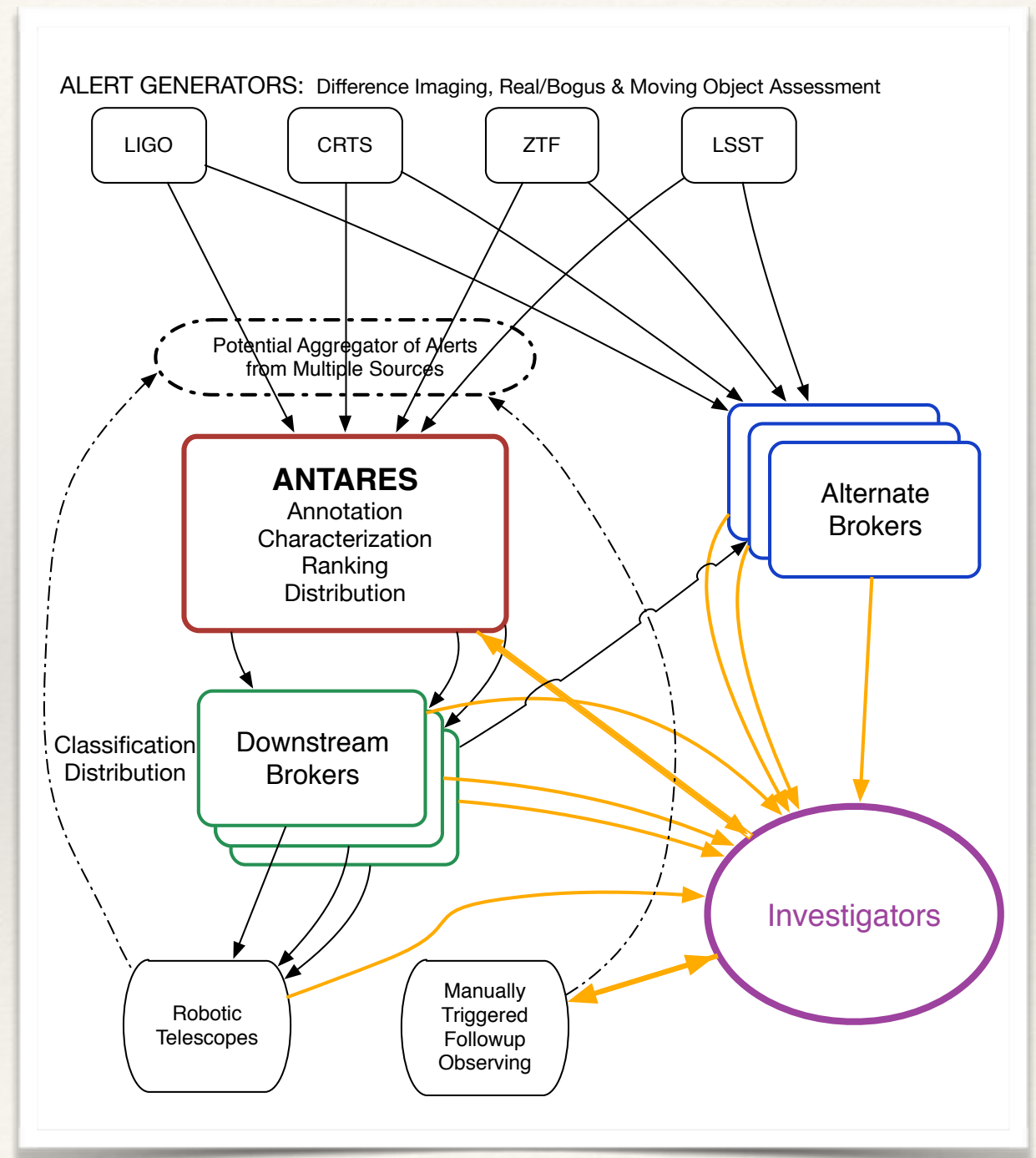
- ❖ LSST cadence (not finalized) strongly affects what we can find
- ❖ Cross-matching is slow & external catalogs change
- ❖ Different “features” are relevant for different classes of variables and transients
- ❖ Many well-studied algorithms aren’t directly applicable

A big problem, with a side of challenges

- ❖ We don't know what we'll find
- ❖ We'll find it very inefficiently
- ❖ We can't agree if it was worth finding in the first place
- ❖ We aren't sure how to find it at all

ANTARES

- ❖ **A**rizona-**N**OAO **T**emporal **A**nalysis and **R**esponse to **E**vents **S**ystem
- ❖ Will winnow down the number of LSST alerts and let you pick the “most interesting”
- ❖ Will serve generic needs - we don’t just care about SNe
- ❖ Archived alerts with contextual information



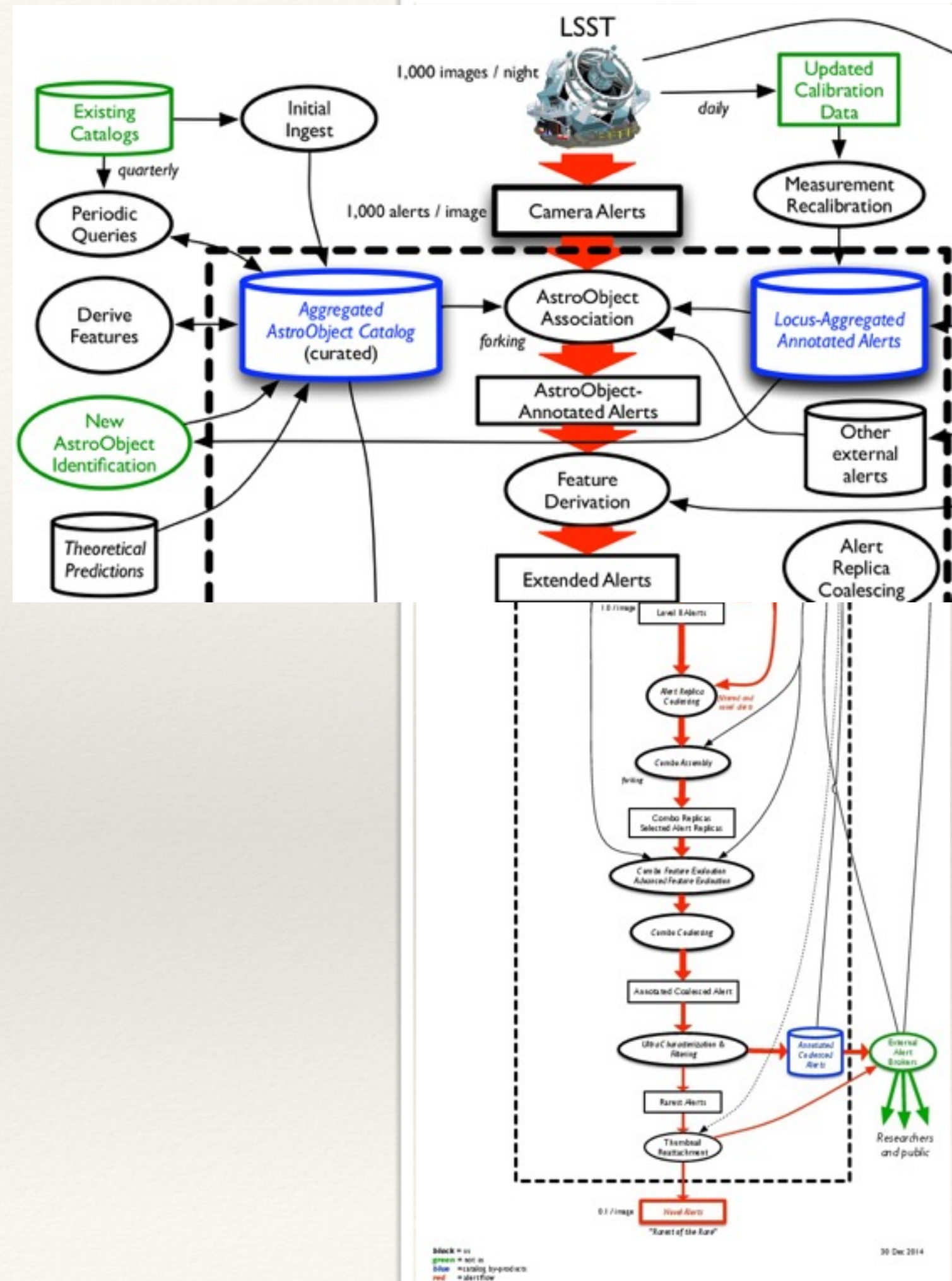
Targets and Timescales

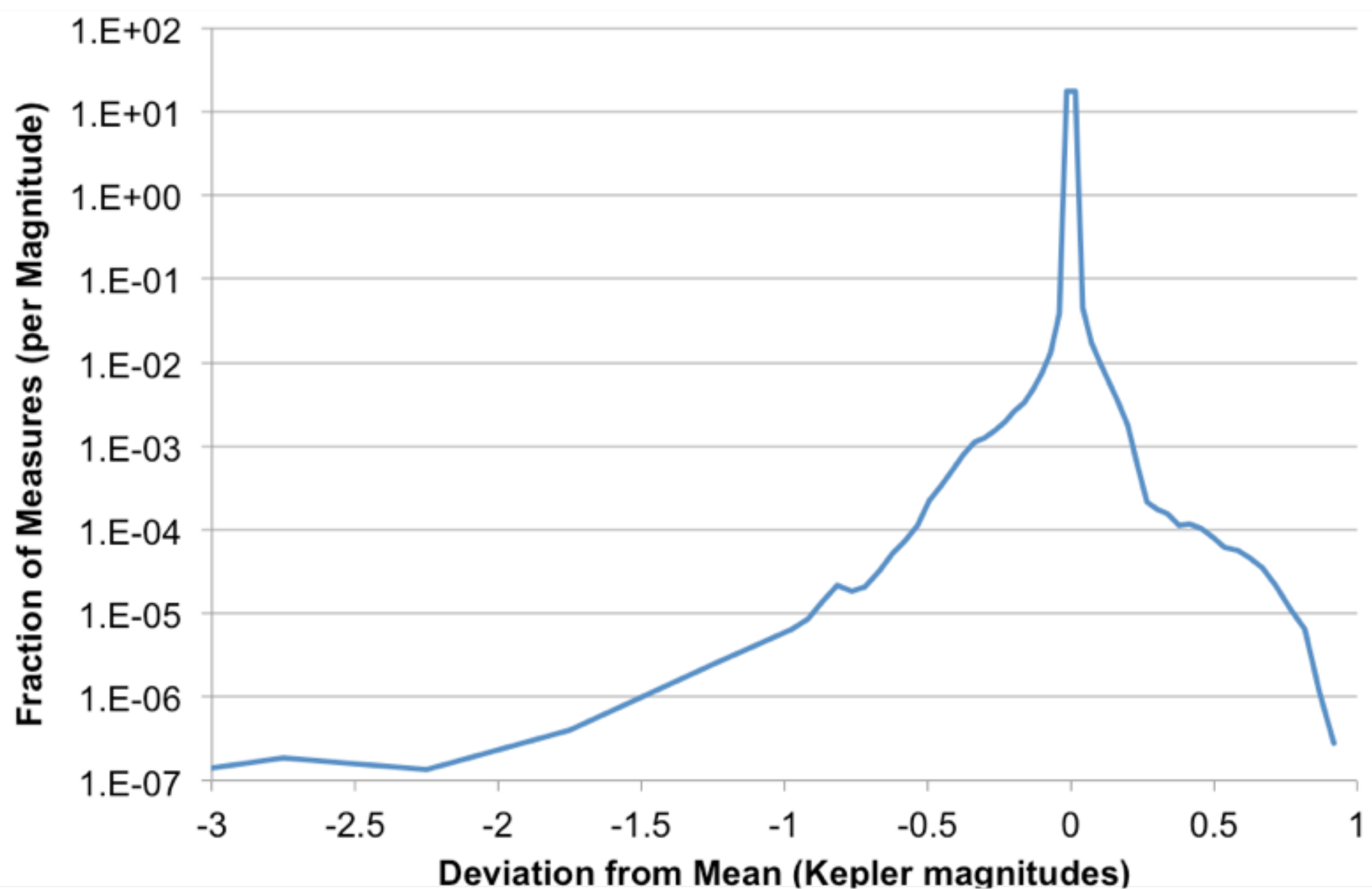
- ❖ **Minutes:** Flares, transits, exotic microlensing, dwarf novae
- ❖ **Hours:** NS-NS mergers, NS-BH mergers, orphan afterglows, stellar rotation, shock breakout, eclipsing binaries, exoplanet phase curves
- ❖ **Days:** Exploding asteroids, WD-WD mergers, WD-NS mergers
- ❖ **Weeks:** Variables and Transients
- ❖ Anything else you might like to find in LSST and follow with current and next-generation ELTs

What we are doing

Saha+ (arXiv:1409.0056)

- ❖ Ingest all LSST alerts, annotate
- ❖ Correlate with other sources (e.g. LIGO)





Annotating Alerts: Adding Value

Can predict variability
given galaxy model and
galactic coordinates of alerts
using Kepler Q13 data
(*Ridgway*+ *arXiv:1409.3265*)

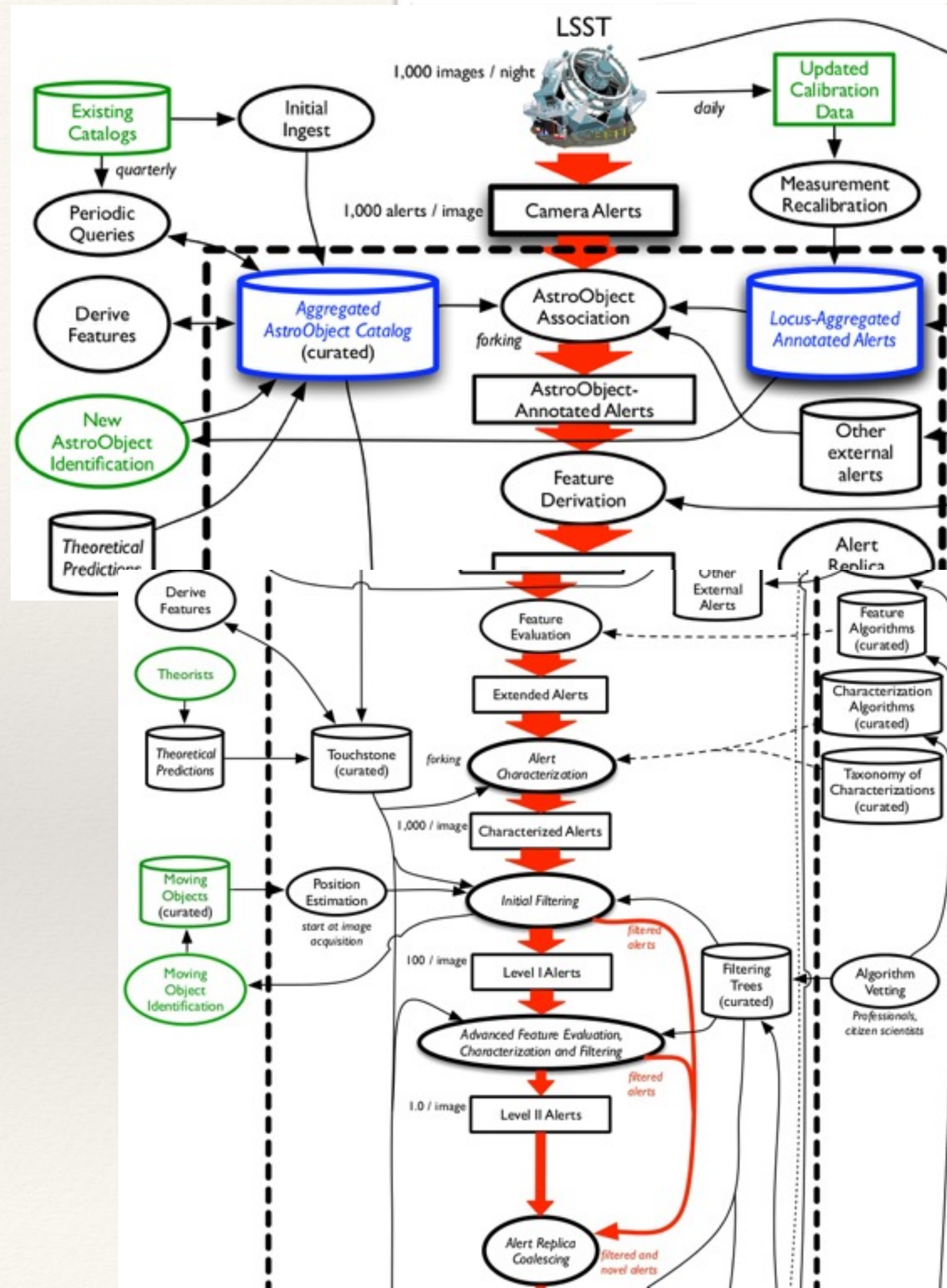
Annotating Alerts: Adding Value

- ❖ Working on collecting and cross-indexing several hundreds of astronomical catalogs (*G. Narayan, T. Matheson, A. Saha, R. Seaman, T. Jenness*) - must be local, will be open
- ❖ Cone search on HTM using SciSQL (*Zhe Wang, Richard Snodgrass*) with adaptation for galactic plane
- ❖ Large Survey Database (Mario Juric, Eddie Schlafly +)
<http://research.majuric.org/trac/wiki/LargeSurveyDatabase>

What we are doing

Saha+ (arXiv:1409.0056)

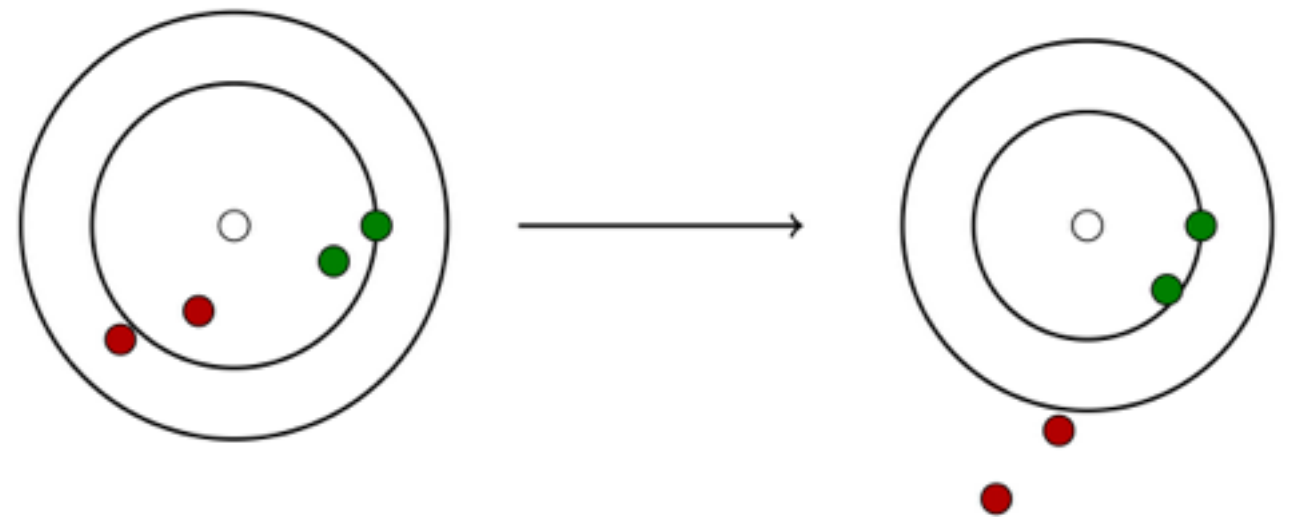
- ❖ Ingest all LSST alerts, annotate
- ❖ Correlate with other sources (e.g. LIGO)
- ❖ Derive features from alerts + annotations + past history
- ❖ Use features to characterize, filter “known sources”



Better Algorithms

- ❖ UA computer scientists developing new algorithms
- ❖ Avoid sensitivity to population size, tolerant of missing data, incorporate uncertainties, work with smaller heterogeneous training sets (*John Kececioglu, Jackson Toeniskoetter + NOAO*)
- ❖ Currently testing on LINEAR, SDSS, and PS1 MDS transients

$$d(p, i) - d(p, t) \geq 1$$



$$\text{error}(p, t, i) = \max\{0, 1 - d(p, i) + d(p, t)\}$$

Better Algorithms

- ❖ UA computer scientists developing new algorithms
- ❖ Avoid sensitivity to population size, tolerant of missing data, incorporate uncertainties, work with smaller heterogeneous training sets (*John Kececioglu, Jackson Toeniskoetter + NOAO*)
- ❖ Currently testing on LINEAR, SDSS, and PS1 MDS transients

$$\sum_p \max_{i,t} \{\text{error}(p, i, t)\}.$$

Maximize distance to nearest impostor
(WCE)

or

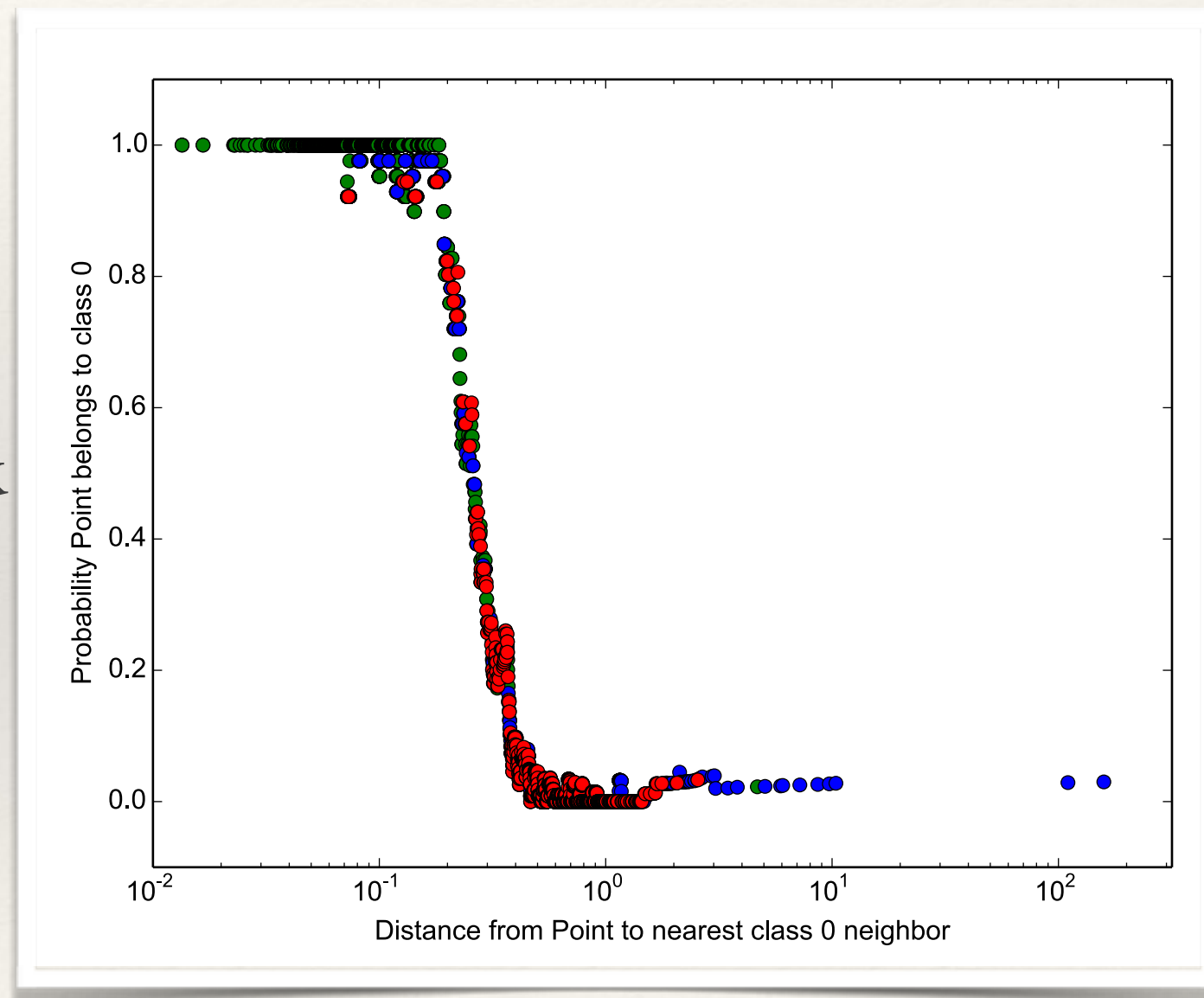
Maximize average distance to impostors
(AE)

$$\sum_{p,i,t} \text{error}(p, i, t).$$

$$\text{error}(p, t, i) = \max\{0, 1 - d_{\ell(i)}(p, i) + d_{\ell(p)}(p, t)\}$$

Better Algorithms

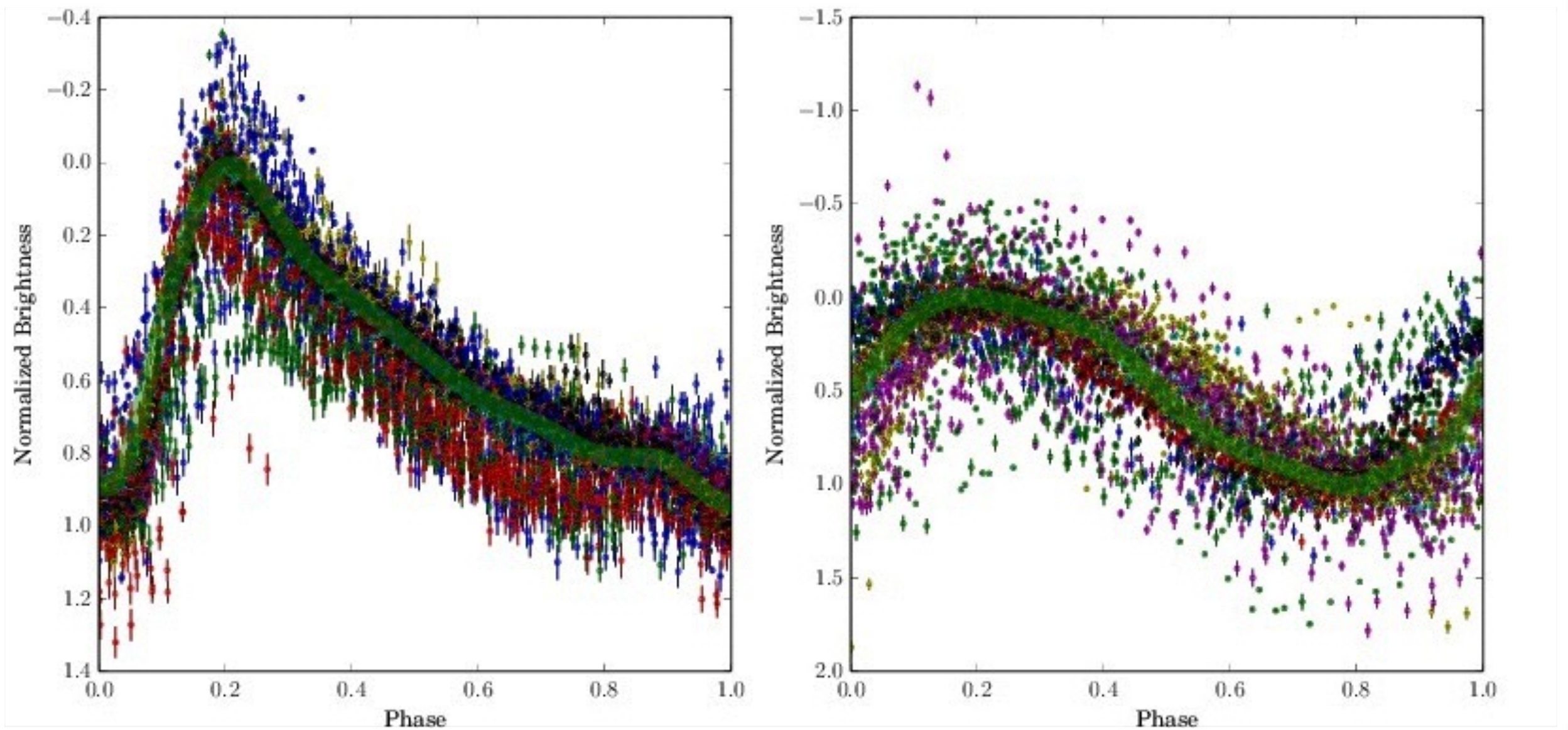
- ❖ UA computer scientists developing new algorithms
- ❖ Avoid sensitivity to population size, tolerant of missing data, incorporate uncertainties, work with smaller heterogeneous training sets (*John Kececioglu, Jackson Toeniskoetter + NOAO*)
- ❖ Currently testing on LINEAR, SDSS, and PS1 MDS transients



Classification Results

- ❖ 10 fold validation on multi-label sets
- ❖ This is a supervised algorithm that needs labelled data
- ❖ Does what observers do - compares candidates to what we already know in complex multi-dimensional feature space
- ❖ **Fast**

	Class 0	Class 1	Class 2	Class 3	Class 4	Average	Overall
Euclidean, kNN	97.02	69.18	88.79	87.11	100.0	88.42	85.52
Euclidean, Emp.	98.53	85.45	76.16	62.75	98.13	84.20	88.75
AE, Equal Weight, kNN	93.94	13.38	95.05	85.71	100	77.62	65.35
AE, Equal Weight, Emp.	98.43	25.74	94.44	63.87	100	76.50	70.41
Matrix, kNN	77.32	73.08	90.91	98.88	98.21	87.68	79.30
Matrix, Emp.	90.39	92.20	71.52	81.79	29.46	73.07	86.80



Marcus Lee (TOCC, NOAO Summer 2014 REU Program, now UA)

Modeling the variable and transient sky

- ❖ Building models of several classes of variables and transients
- ❖ Simulations to ANTARES end-to-end, *ultra-characterization*
- ❖ Similar architecture to SNANA for SNe



Combos

Associating an alert with individual objects and forking may not capture all the information

Combos coalesce forked alerts to find events that are interesting with a wider context

Progress Report

- ❖ **Prototype construction in full swing (v1 Summer 2015)**
- ❖ Scaling and optimization through 2016
- ❖ Run on live alert streams in 2017
- ❖ Likely NCSA collaboration during LSST operations
- ❖ LSST and ANTARES can use input from **You**

If you are interested in rare & exotic transients,
we'd like to know what they look like.
We can help you find them.