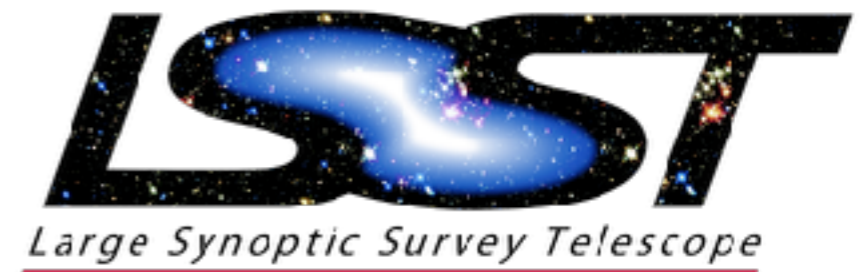


# Optical transient surveys of today and tomorrow : machine learning applications

Stephen Smartt  
Queen's University Belfast

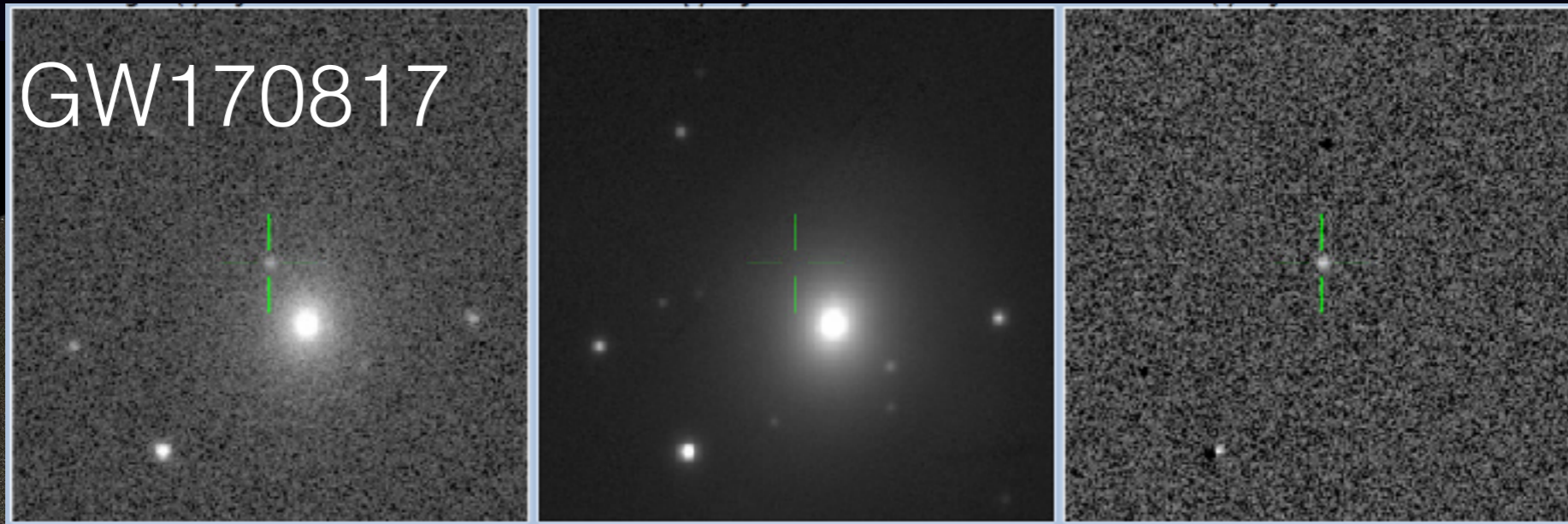


Ken Smith, Dave Young  
Darryl Wright  
Amanda Ibsen

ZOONIVERSE



# The Pan-STARRS Sky



*This is a gri colour image of the  $3\pi$  Steradian survey. Image quality is  $\sim 1$  arcsec, with  $0.256''$  sampling over 30,000 square degrees or about 6 Petapixels ( $10^{12}$ ) with over 100 epochs.*

**ATLAS sky : 50 sq degree survey facility in Hawaii -  
all sky, every 2 nights**



# TRANSIENT NAME SERVER

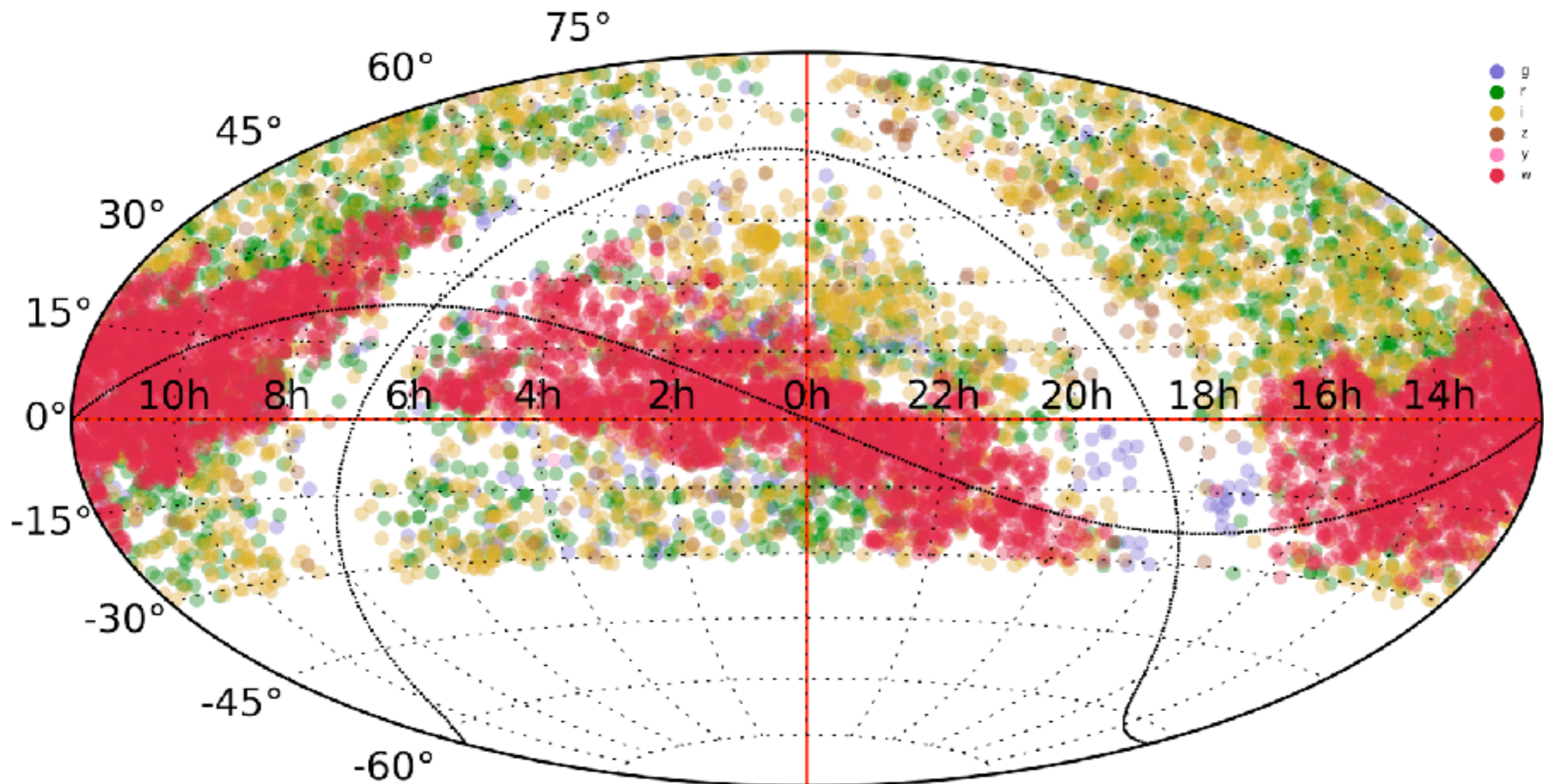
SEARCH

BOTS

LIGO GW



SUPERNOVA  
WORKING GROUP



**PUBLIC transients** reported since Jan 1, 2016

18401

**PUBLIC transients** for the top 5 contributing groups

Pan-STARRS1	8877
GaiaAlerts	4688
ATLAS	2137
ASAS-SN	610
iPTF	550



# Transient sky populations

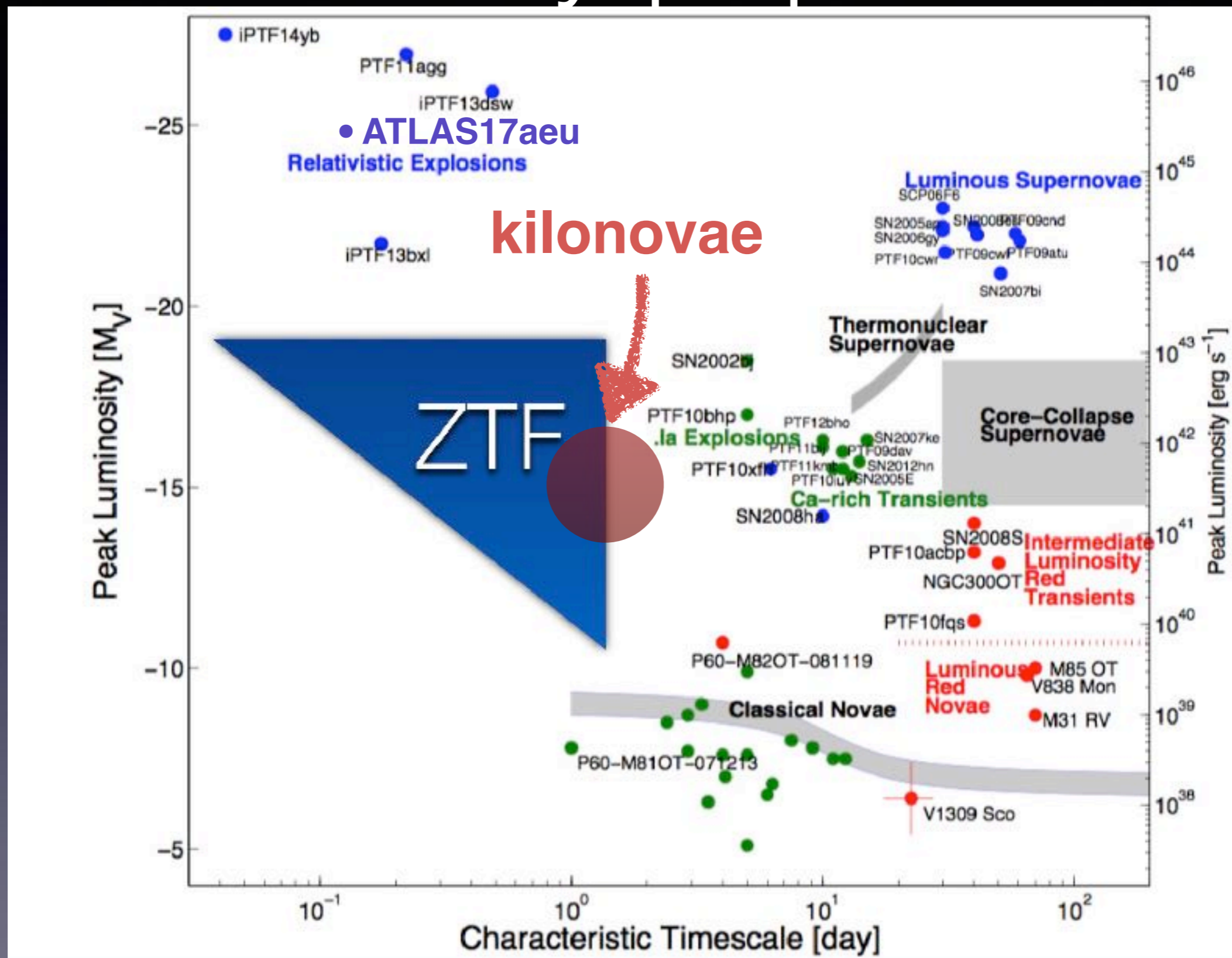


Image Credit : S. Kulkarni, Caltech

# The Universe's degenerates

- White dwarfs



Earth radius, 6371 km

$$m = 1.4M_{\odot}$$

$$\rho = 10^4 \text{ kg/cm}^3$$

- Neutron stars



10km radius

$$m = 1.4 - 2 M_{\odot}$$

$$\rho = 6 \times 10^{11} \text{ kg/cm}^3$$

- Black holes



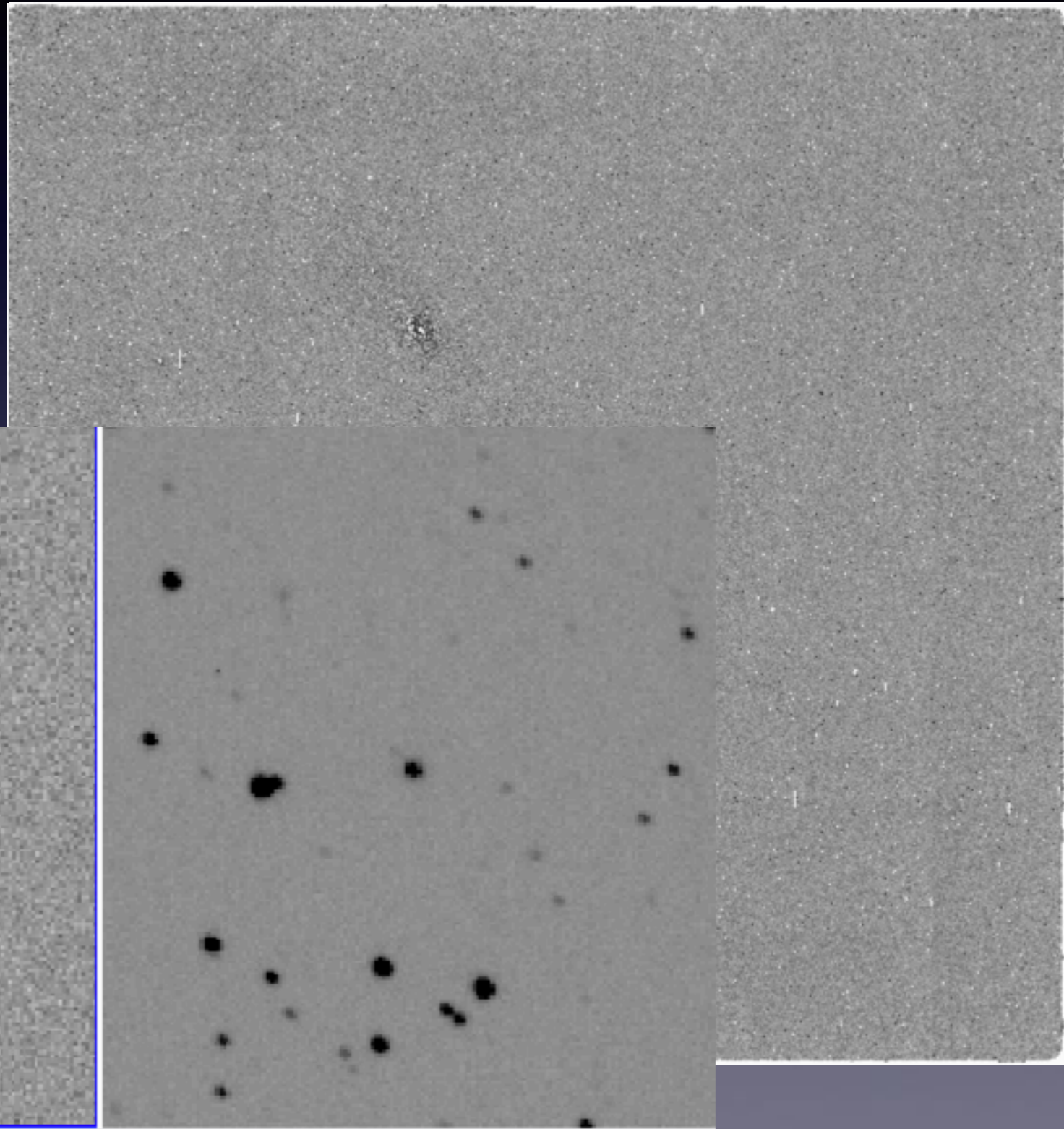
$$m > 2 M_{\odot}$$

R ~ 5km radius

$$R_s = \frac{2GM}{c^2}$$

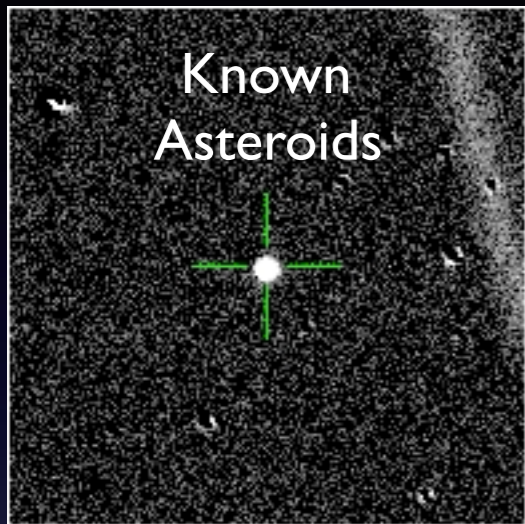
# Typical data frame

- Most common “source” ?
- 2nd most common source (aka “Vermin of the skies”) ?
- 3rd most common source ?

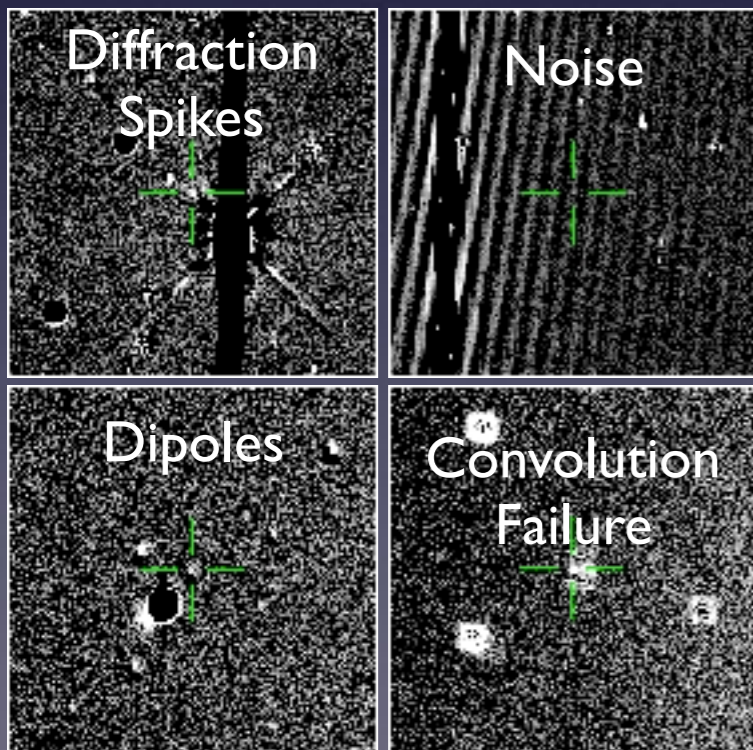


# Machine Learning to junk the bogus

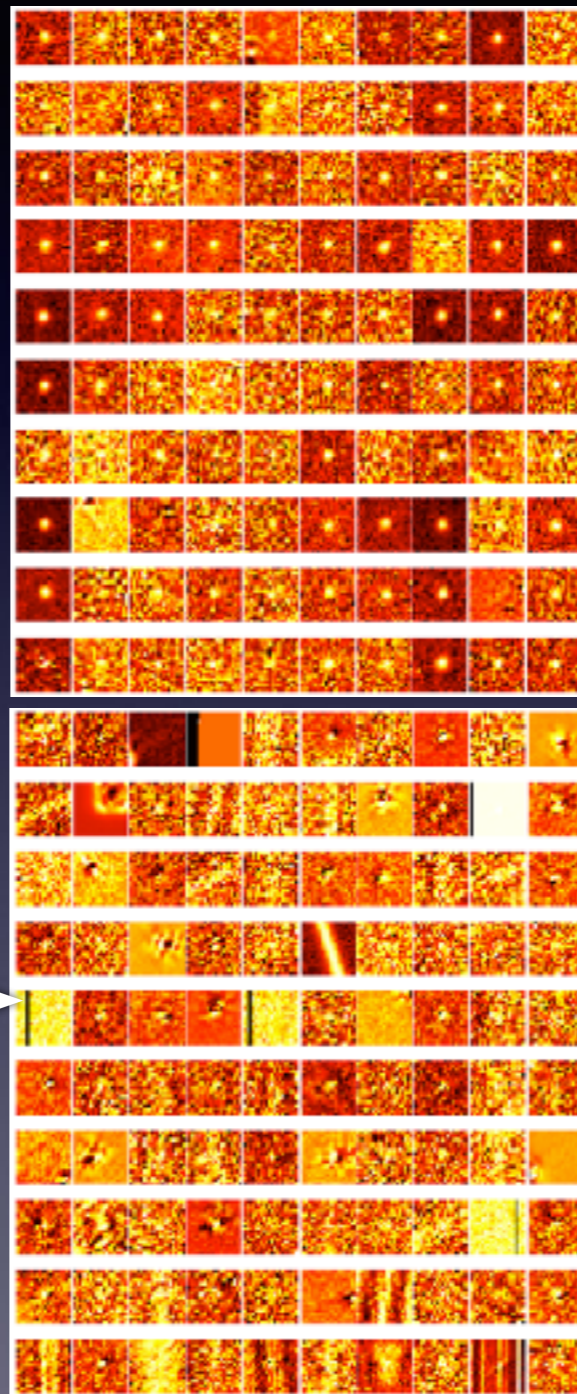
Real Objects



Bogus objects - e.g.:

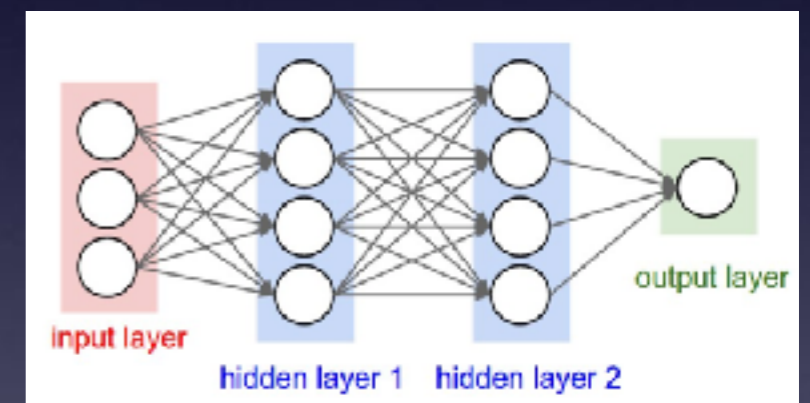


6 ecliptic nights  
250,000 REAL Cutouts



6 ecliptic nights  
750,000 BOGUS Cutouts

Train Classifier

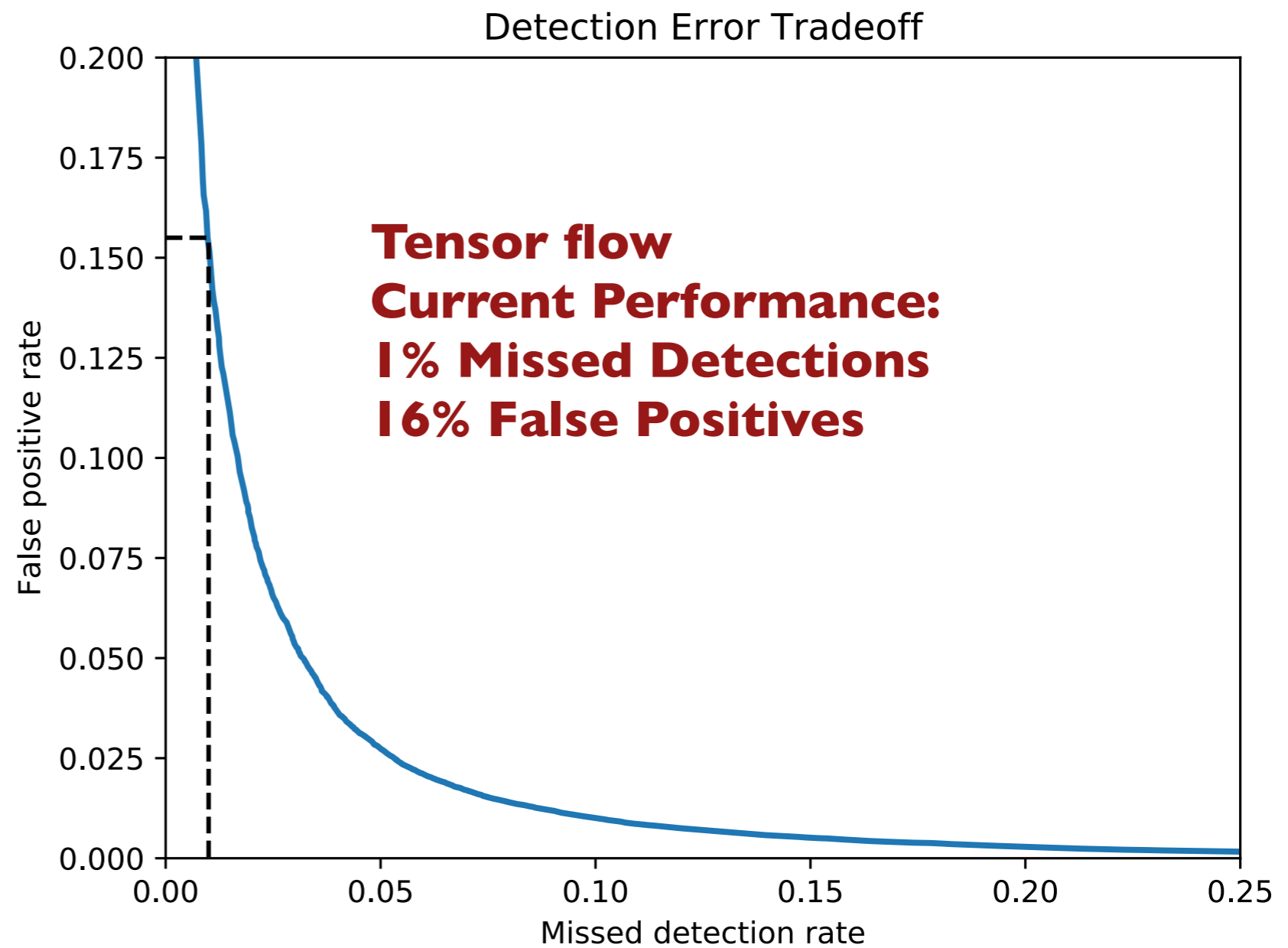
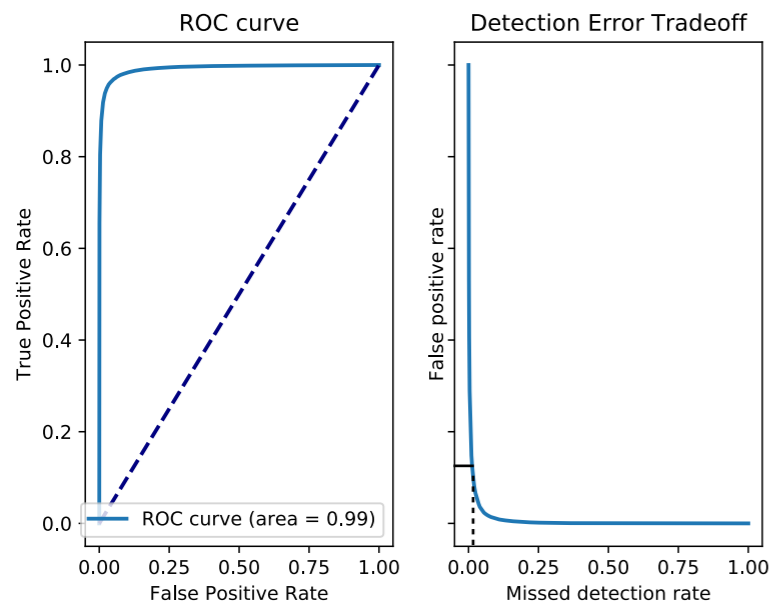
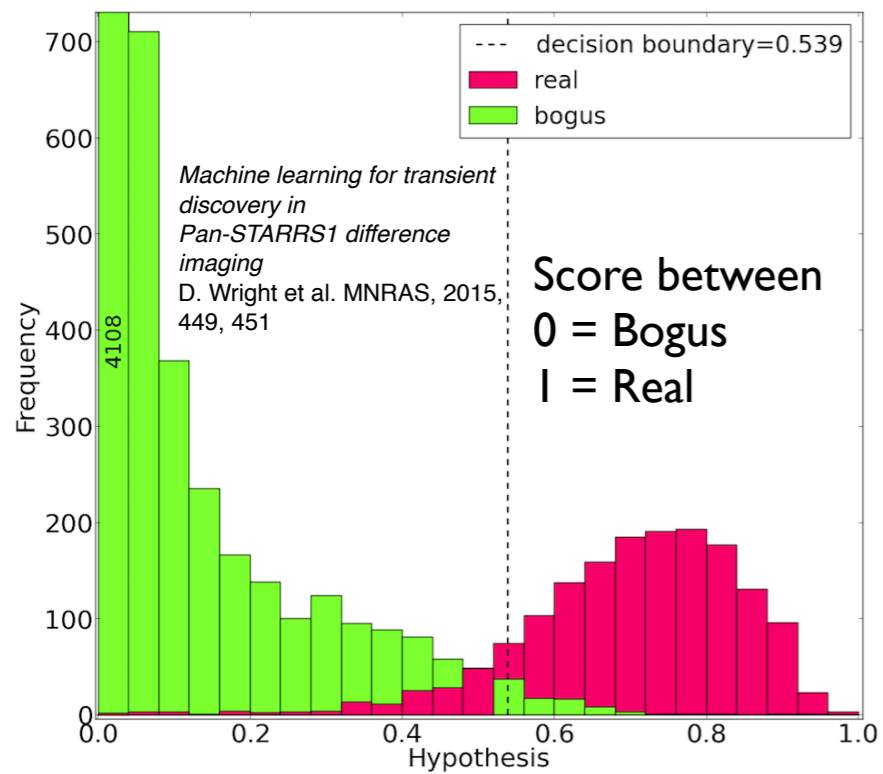


3 layer  
(plus dropout to avoid overfitting)  
Convolutional Neural Network +  
Softmax filter

Trained Classifier applied to hourly incoming stream of objects  
Objects (up to 6 images per object) with median RealBogus Score < Decision Boundary rejected



# Best results to date



# Classification - host object

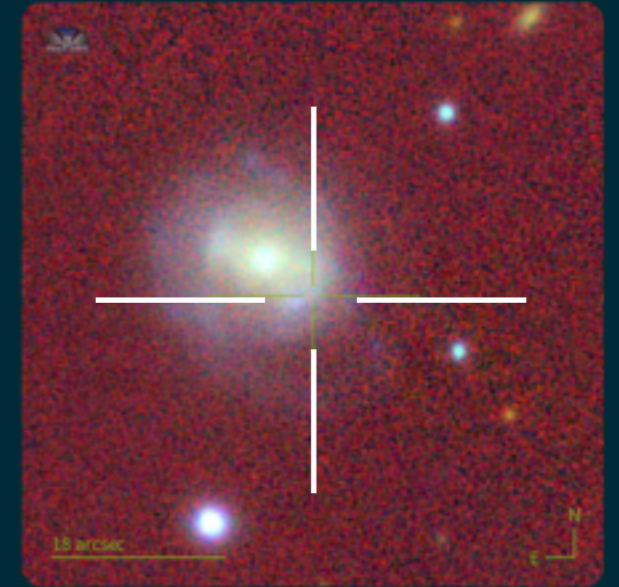
## Star or galaxy and separation

- angular separation crossmatch radius
- physical separation crossmatch radius (if redshift)
- source magnitude filtering
- magnitude dependent search radii for bright stars and galaxies

## Decision Tree Classifications & Reliabilities

- transient given a predicted classification based on the parameters of the catalogued source it matches against
- transients can be given multiple classifications which are later ranked
- a transient can either be **synonymous** with (within 0.5"), **associated** with (>0.5" away) or **annotated** by a catalogued source

host info



[exact sdss location](#)

[sdss nearest object](#)

contextual classification:

**SN** - The transient is possibly associated with **SDSS**

**J161600.57+221608.2**; a 15.60 mag galaxy found in the NED catalogue.

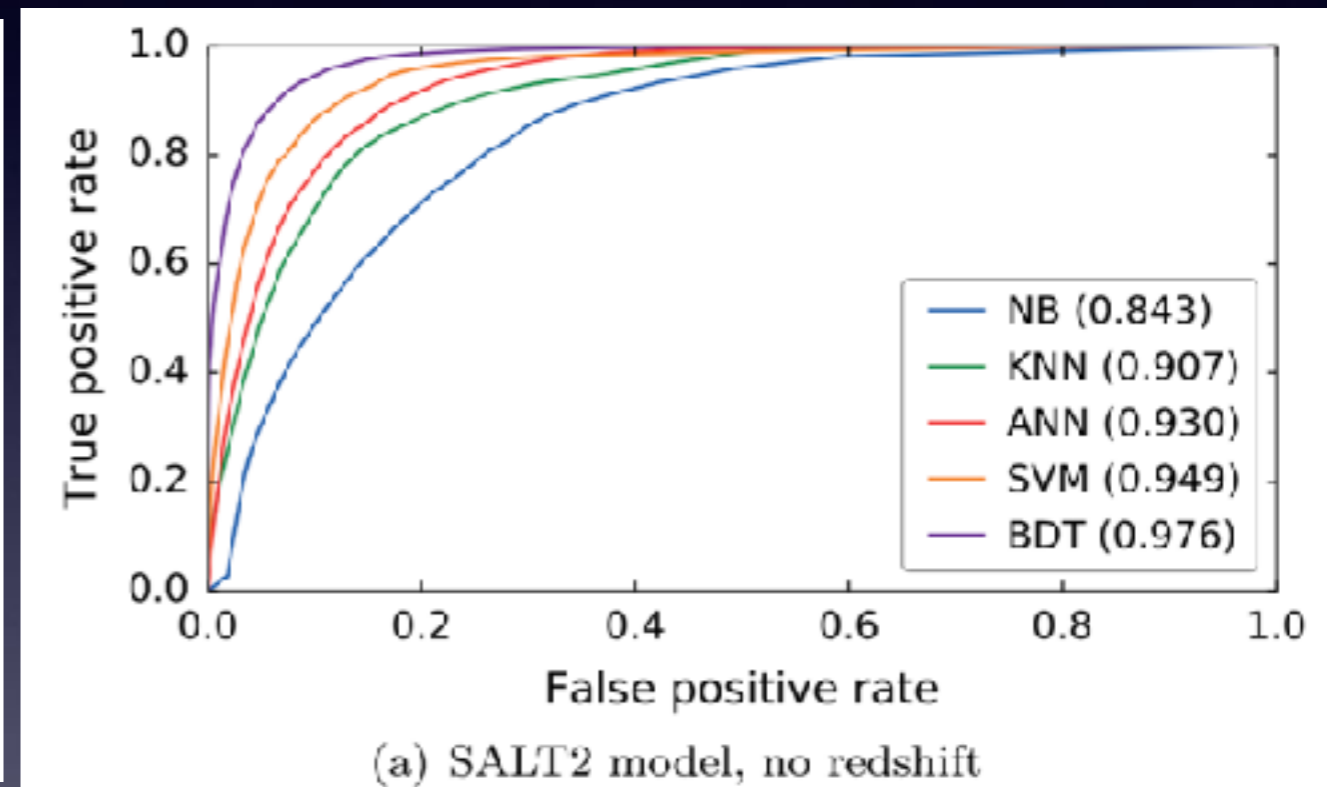
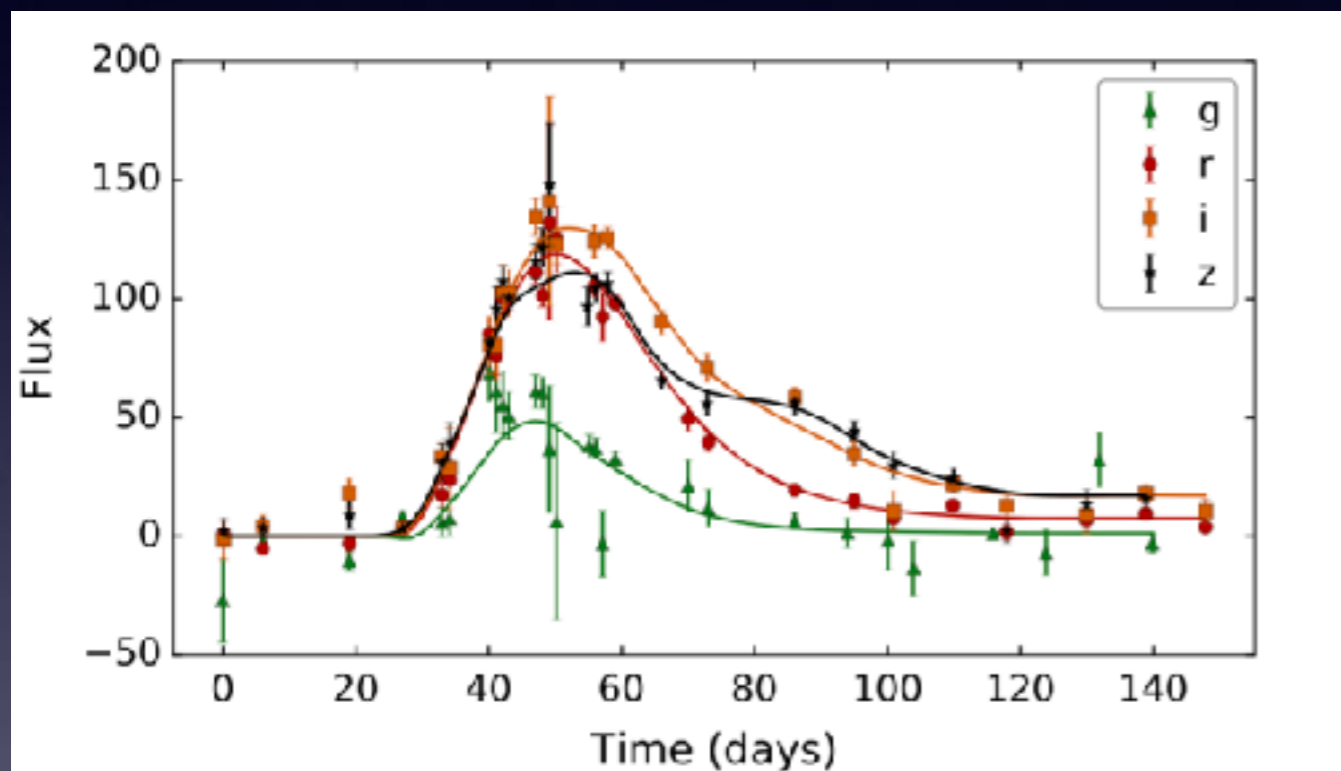
It's located 3.46" S, 4.86" W (1.7 Kpc) from the galaxy centre. A host

$z=0.014$  implies a transient  $M = -20.41$ .

# Lightcurve Classification of transients

Lochner, McEwan, Peiris, Lahav, Winter 2016

Extract features + ML on features



Achieved : 90% pure & 84% complete

Missed detection rate of 1%, false positive rate ~ 20-30%.

# Large Synoptic Survey Telescope

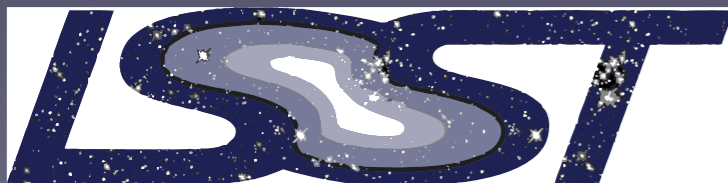


## The Summit Facility



LSST@Europe 3, Lyon, France, June 11-16, 2018

Slide credit :  
Steve Kahn  
LSST@Europe3  
Lyon, June 2018



 LSST:UK Consortium

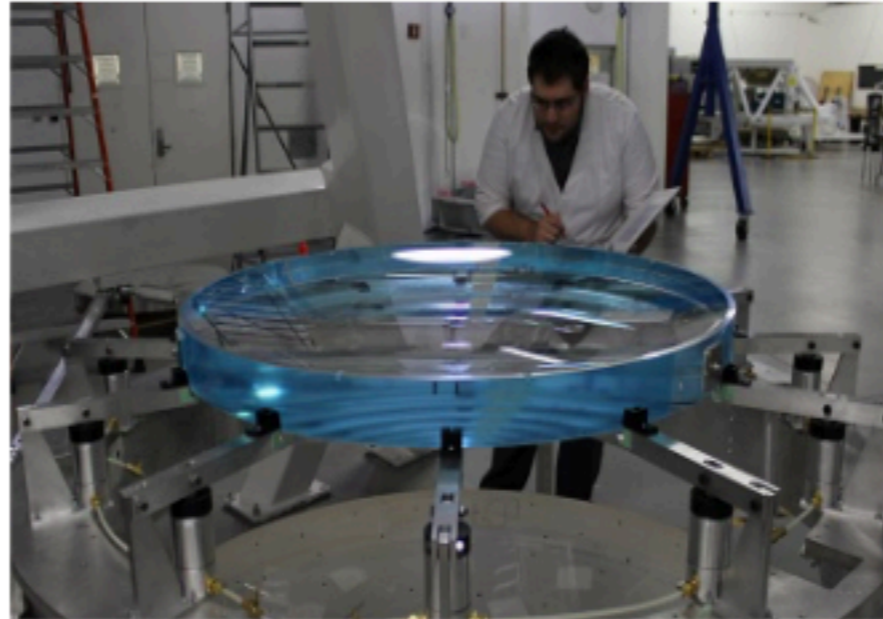
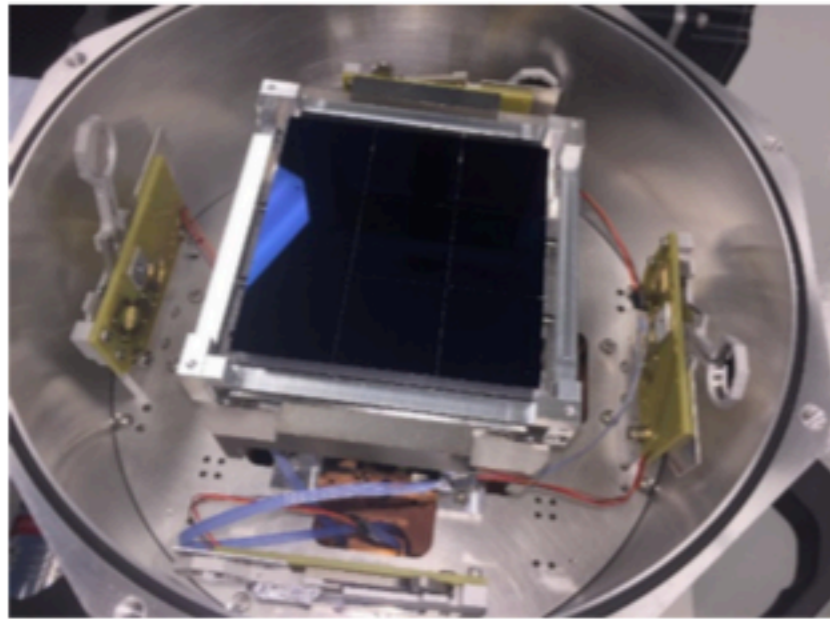


Science & Technology  
Facilities Council

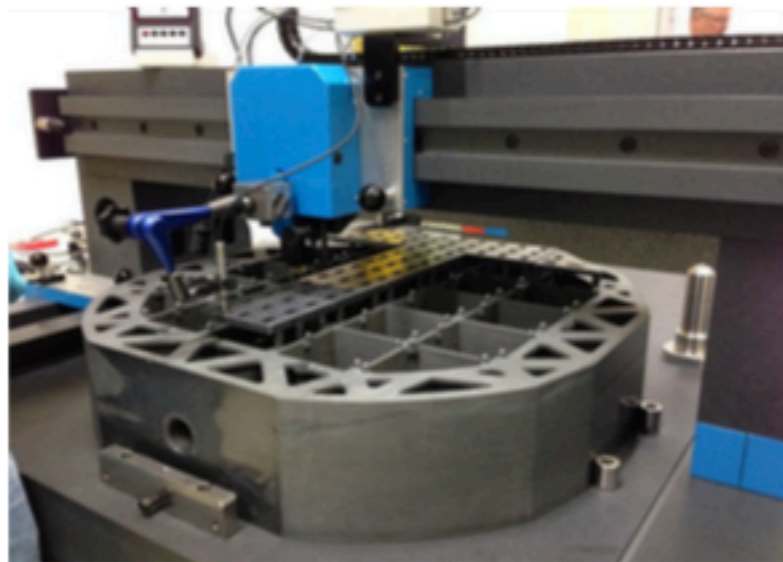


Photo credit :  
Owen McBrien

# Large Synoptic Survey Telescope



Slide credit :  
Steve Kahn  
LSST@Europe3  
Lyon, June 2018



**ATLAS  
Field of View**

**5.4°**

**LSST  
10 square  
degrees**

Pan-STARRS 1



# LSST - $10^7$ alerts per day

Junk

ML

Vermin (asteroid)

spatial coincidence or trail

Star

AGN

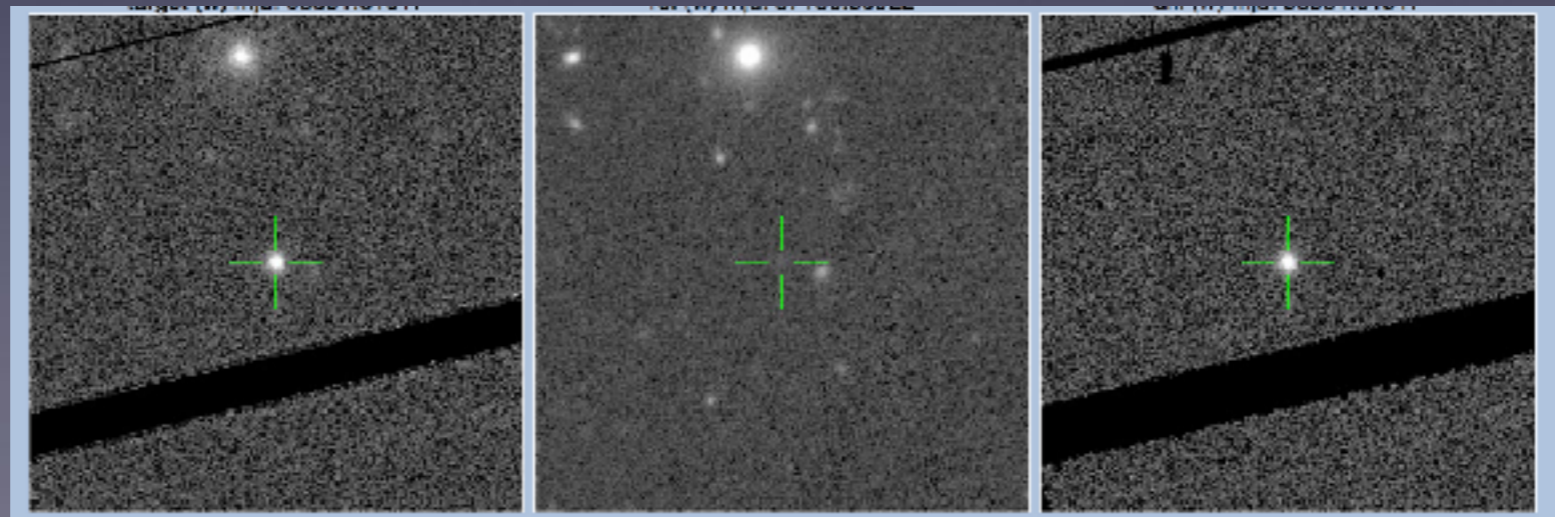
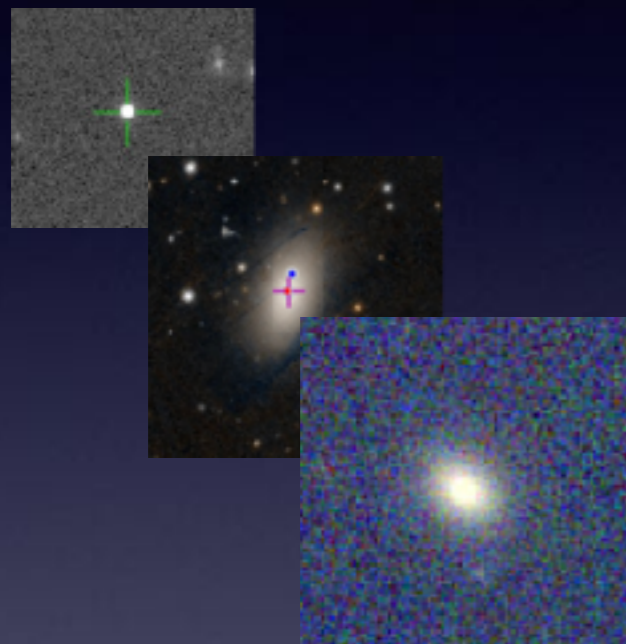
Nuclear Transient

Supernova

orphan

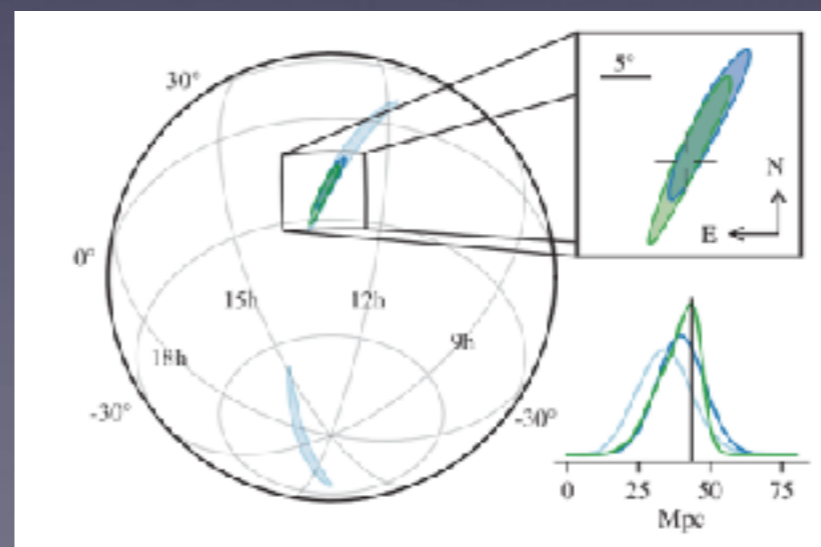
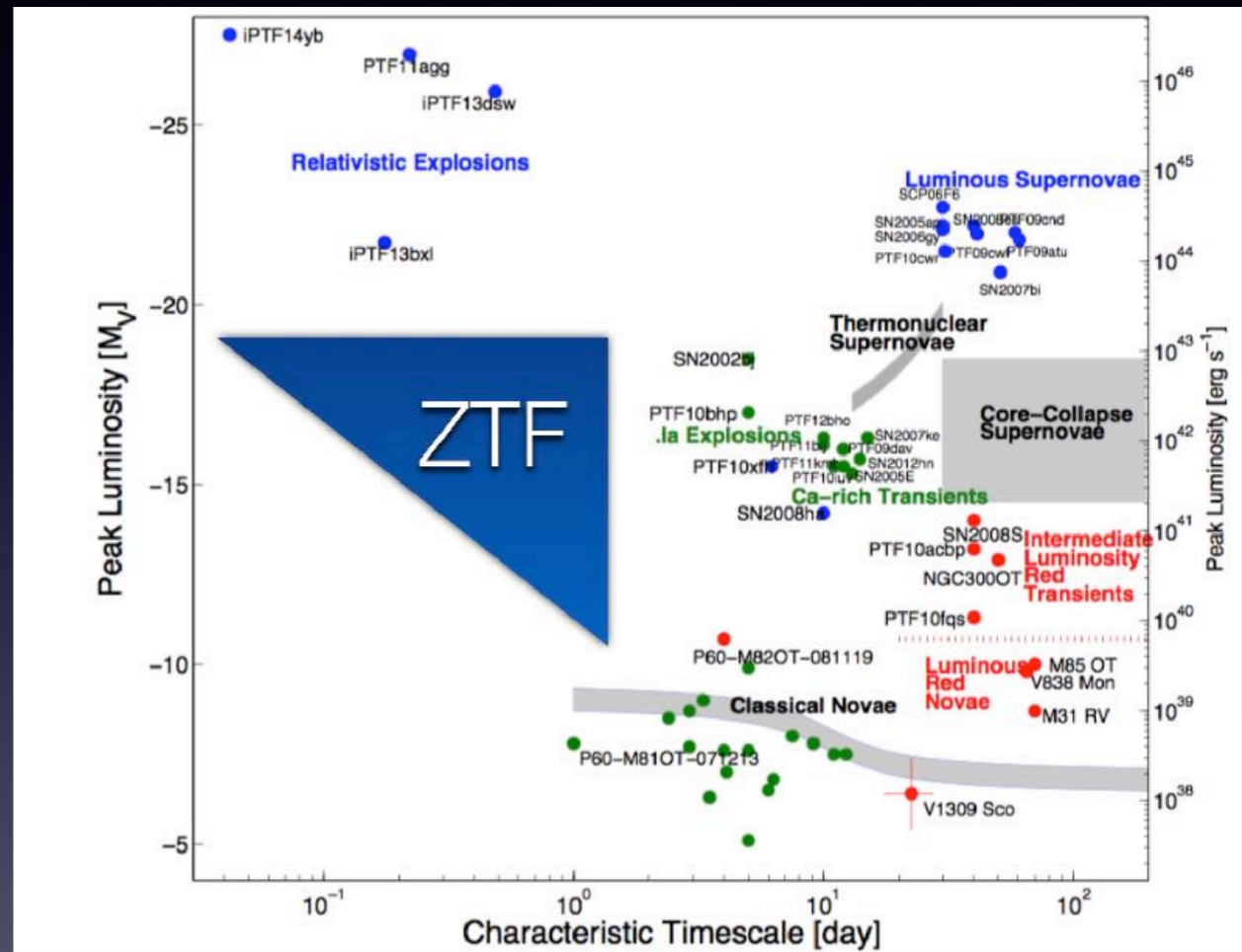


Boosted  
Decision  
Tree



# Summary : scratching the surface of potential for LSST

- Multiple machine learning to provide probability of physical classification
- Identify the outliers
- Real-time, every day - ability to trigger rapid follow-up
- Cross match with radio, x-ray, gamma-ray surveys, LIGO-Virgo sky localisation maps for GW
- Enormous discovery potential in data, but completeness and probabilistic approach essential
- **Understand the population - repeatability and completeness**

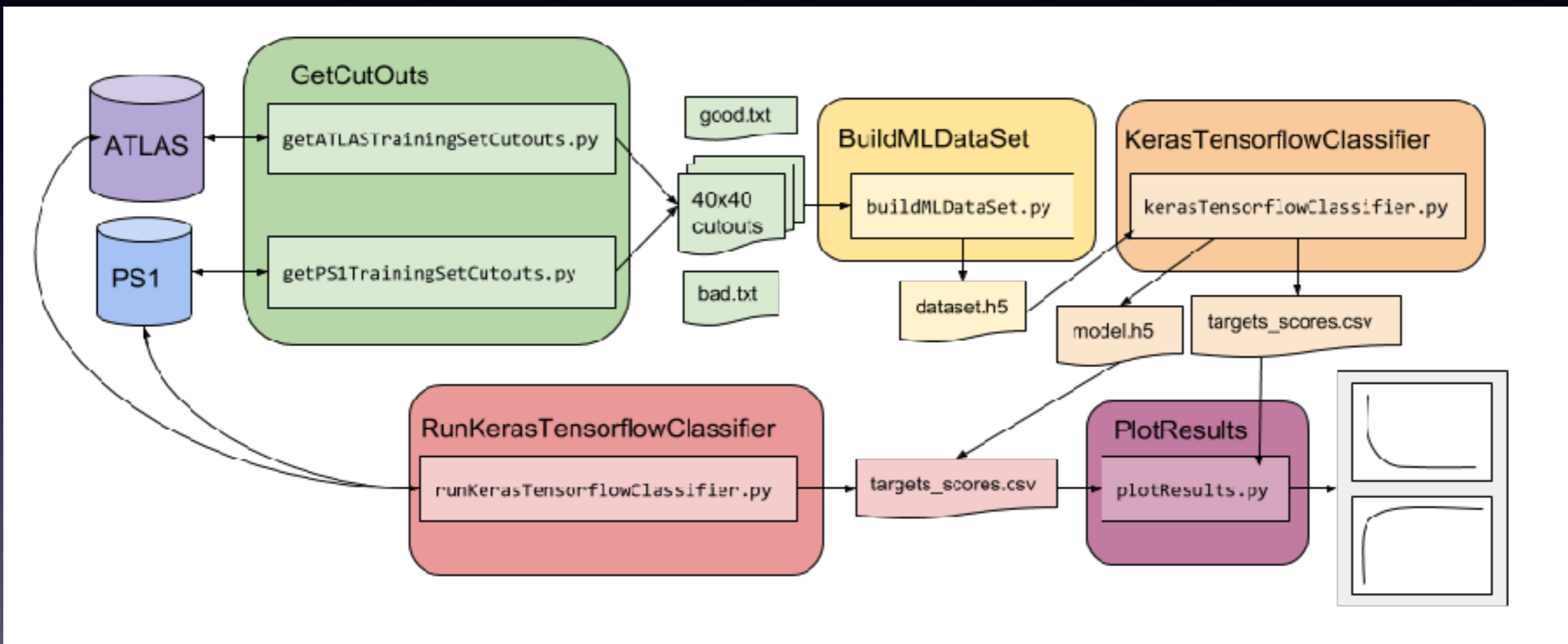


**GW170817**



# ATLAS + Pan-STARRS Machine Learning Trainer Pipeline

Automated End to End Training - User just specifies which nights to be used



Ibsen & Smith - see <https://github.com/aibsen/ATLAS-ML>