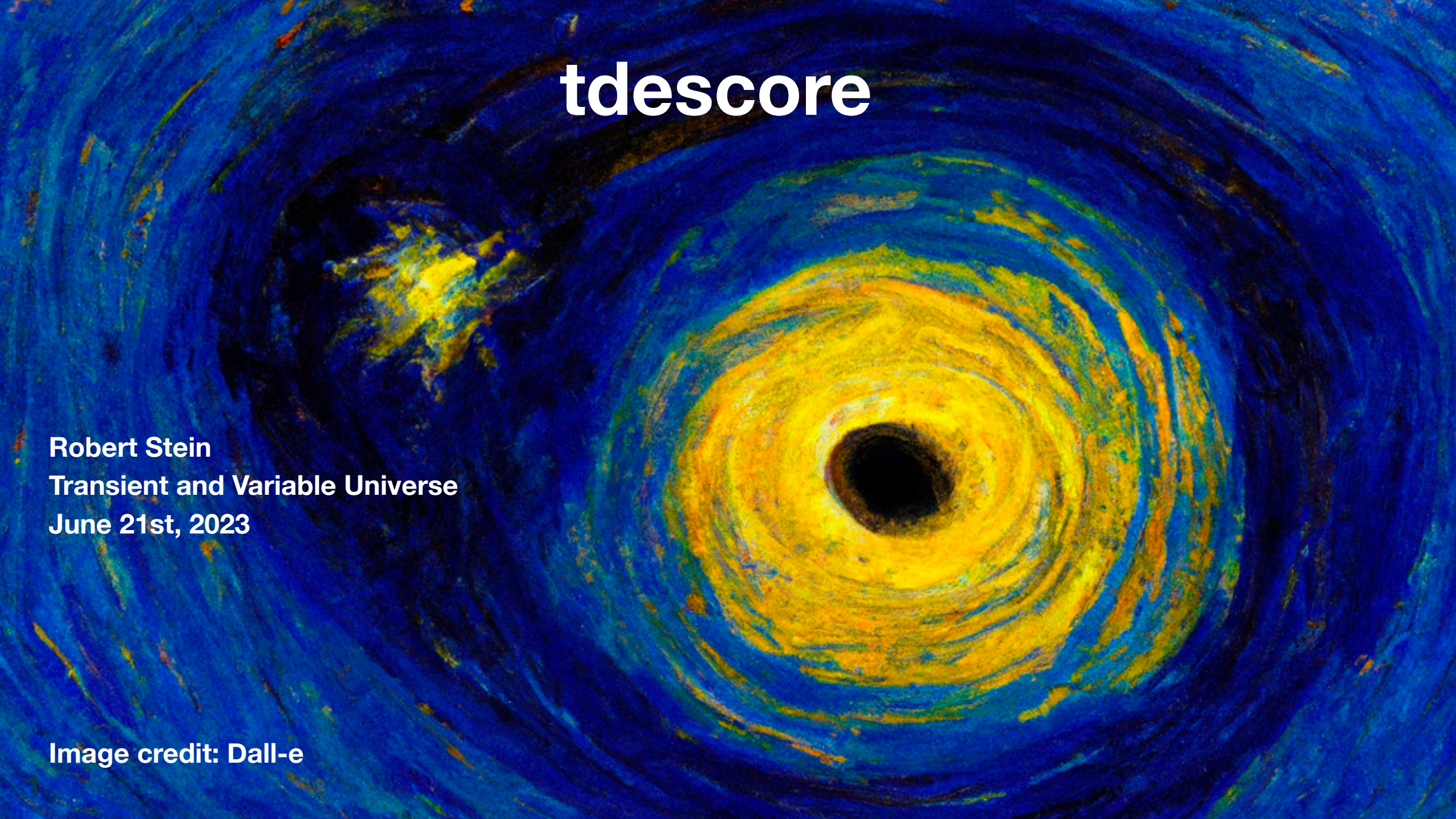
tdescore

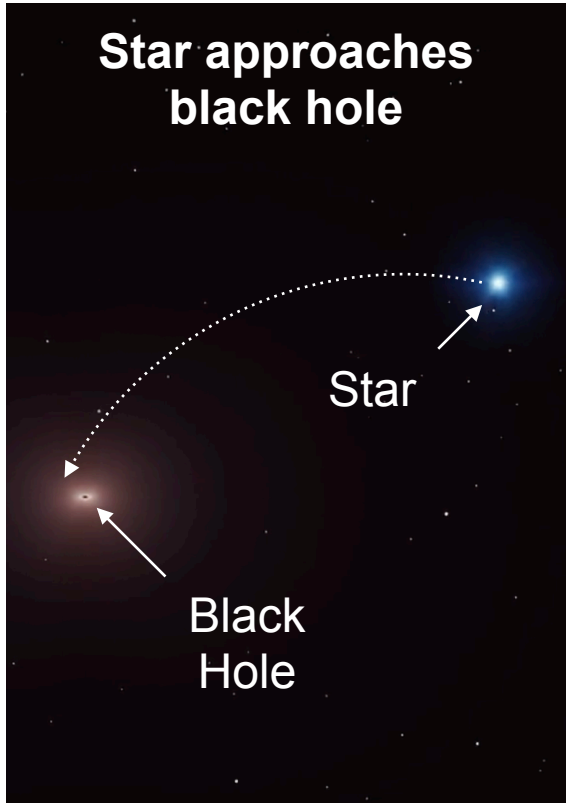


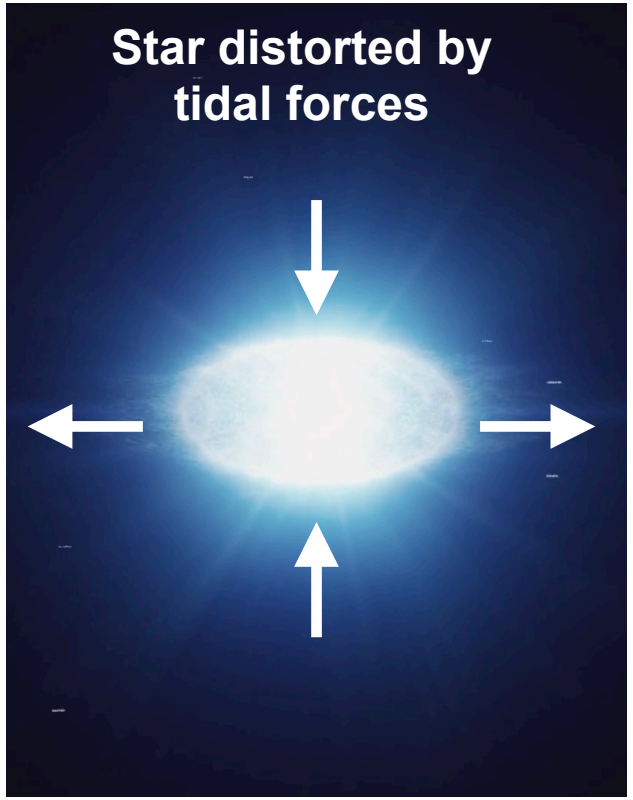
Robert Stein
Transient and Variable Universe
June 21st, 2023

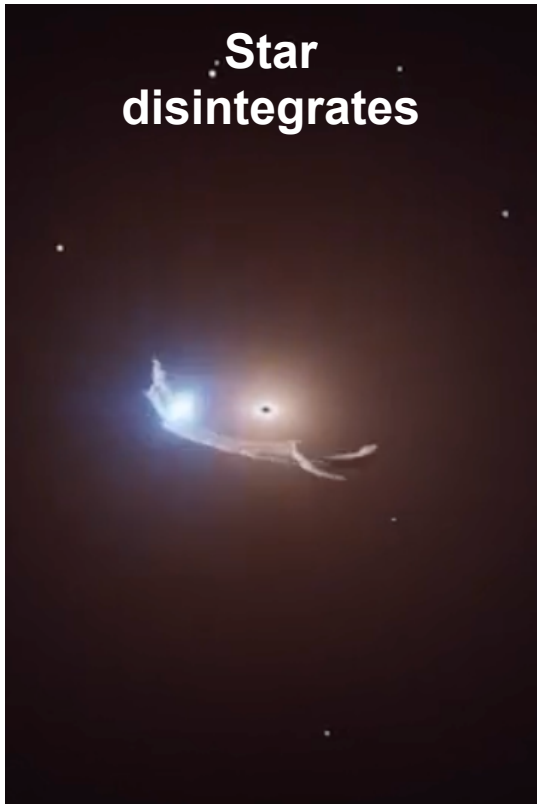
Image credit: Dall-e

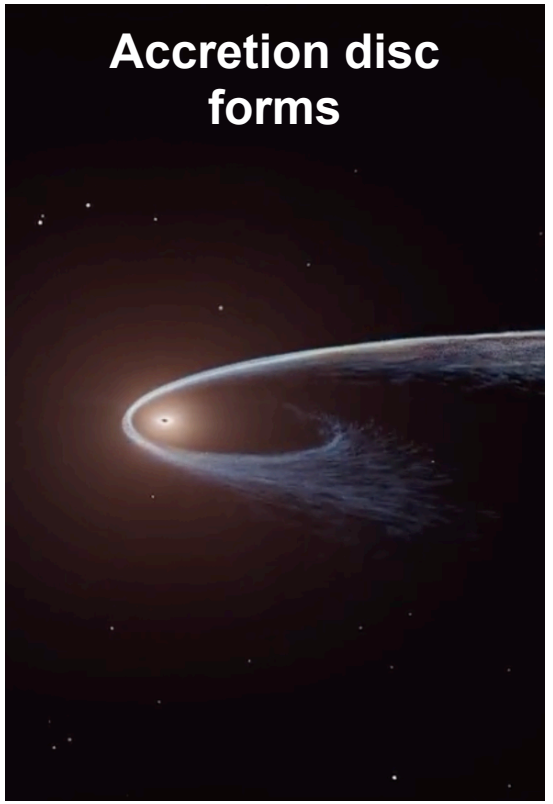
A brief introduction to TDEs

What are TDEs?









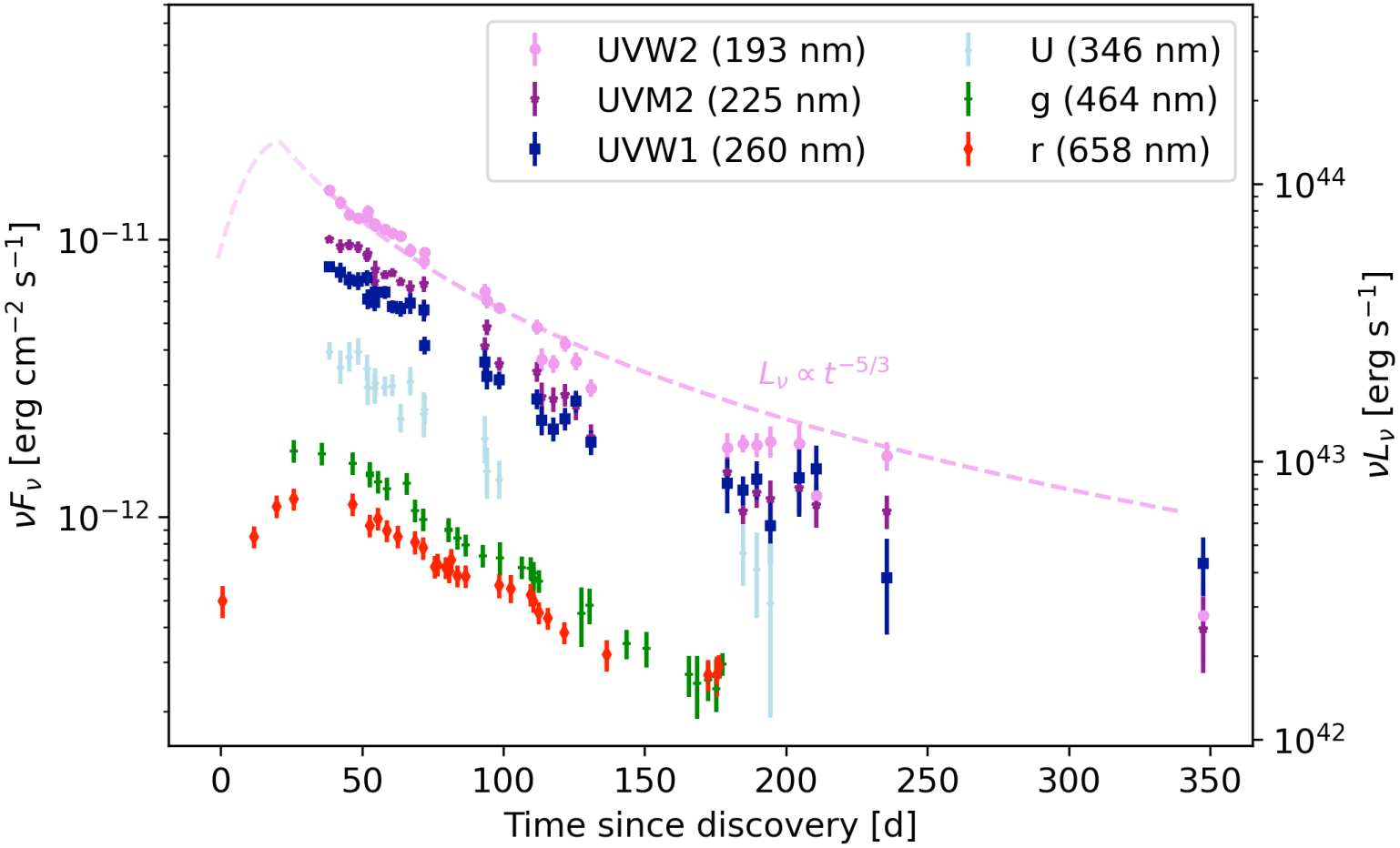
1 2 3 4

Credit:DESY/SciComLab

What are TDEs? - An empirical definition

TDEs are (to first order):

- **nuclear flares**
- **Intrinsically hot/blue**
- **rise over a period of ~weeks**
- **and then fade over months**
- **with little apparent cooling**



The TDE Explosion

1 TDE
per month

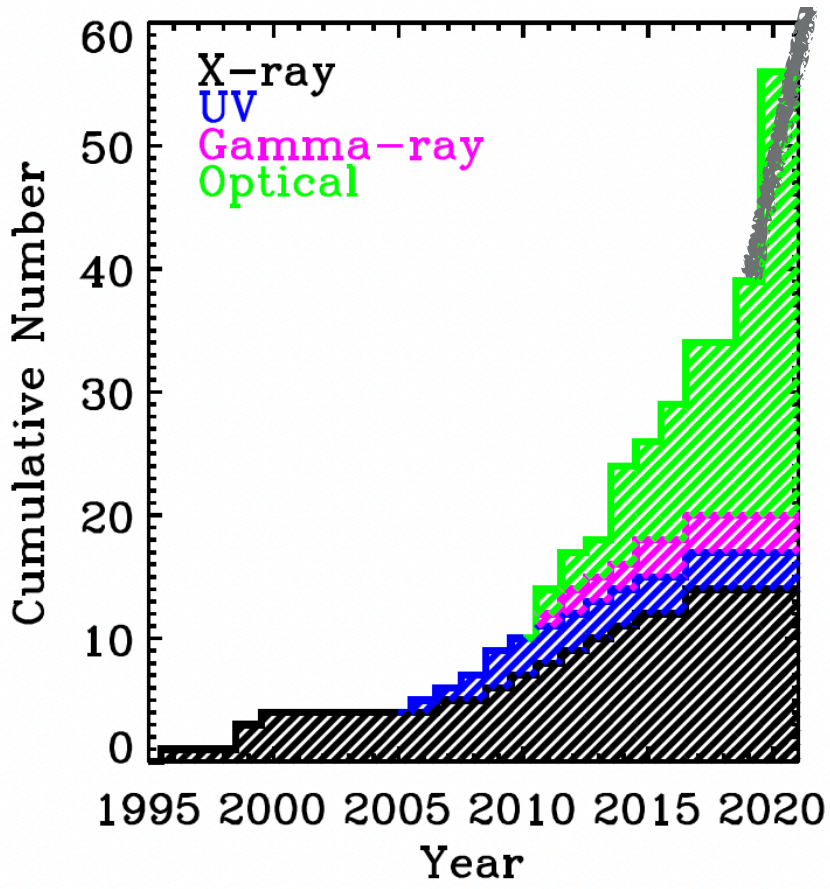
ATLAS

ASAS-SN

Pan-STARRS

Gaia

ZTF



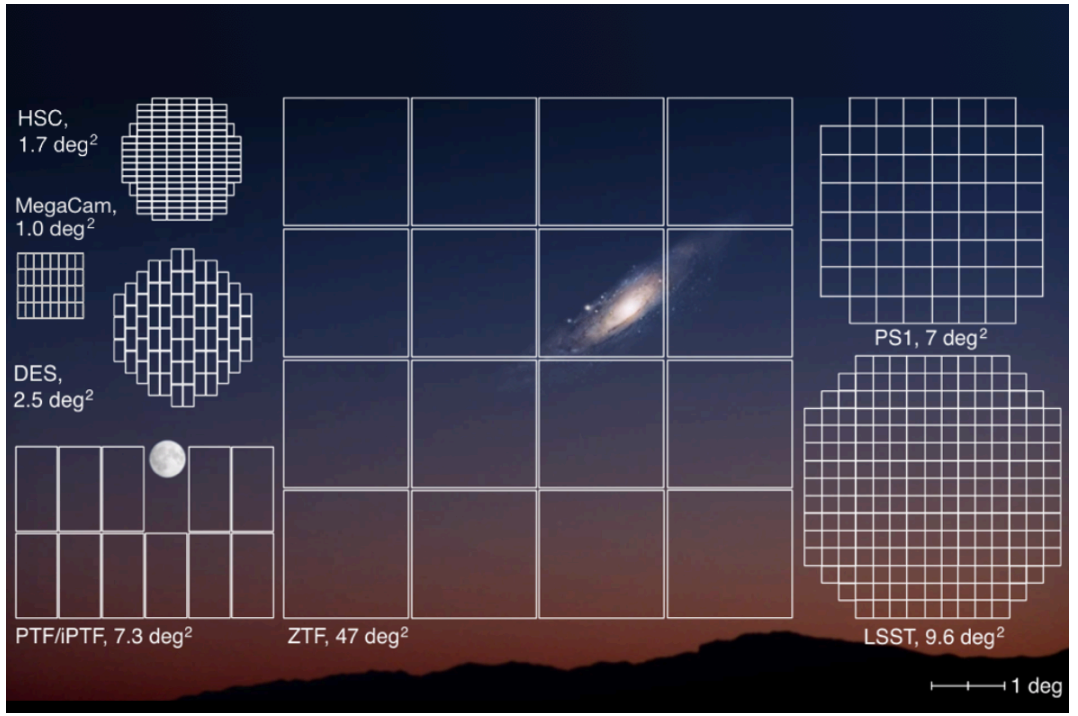
Gezari (2021)
https://doi.org/10.1146/annurev-astro-111720-030029

The ZTF TDE Program

Introducing the Zwicky Transient Facility (ZTF)

Credit: Iair Arcavi

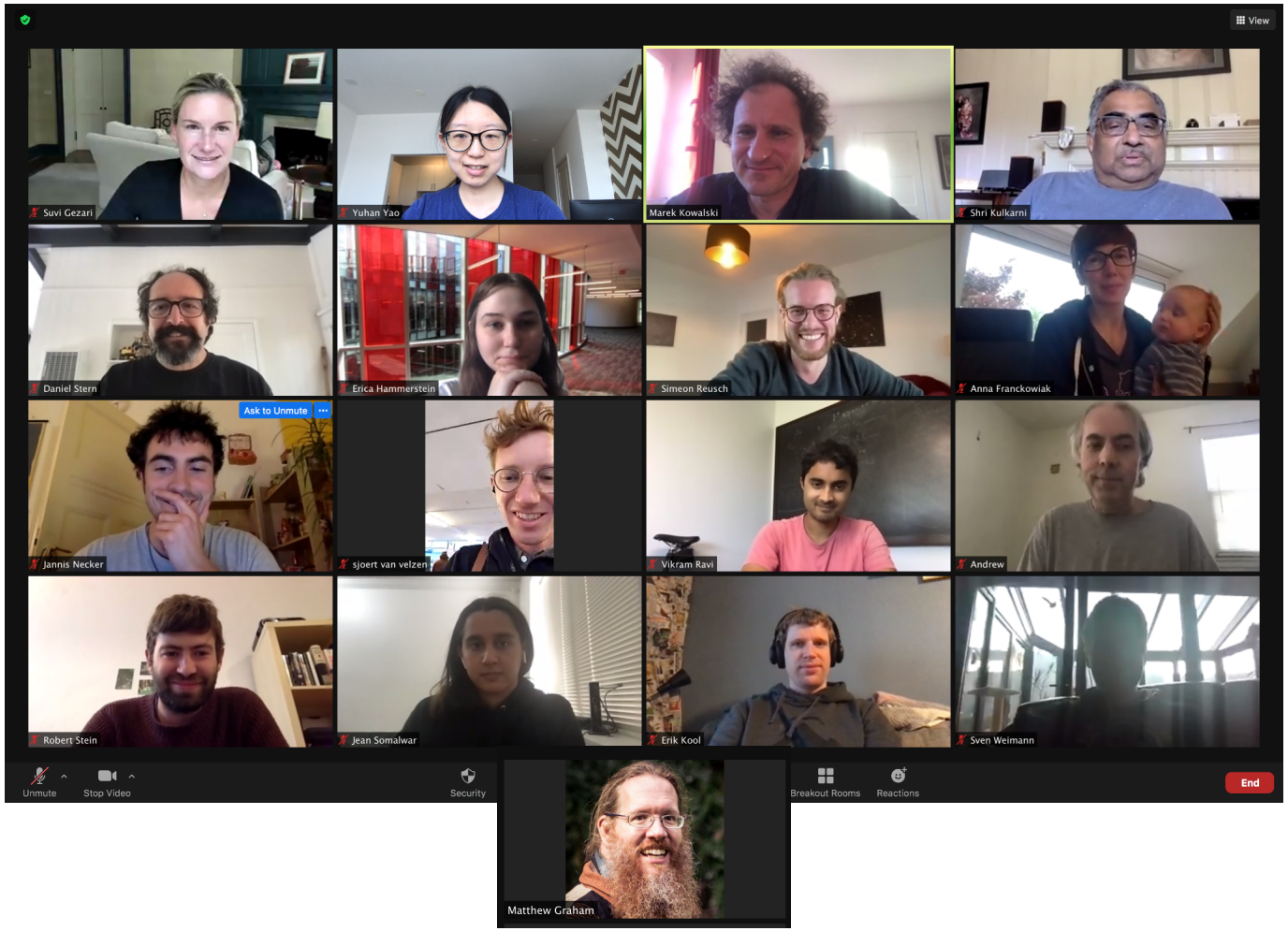
Credit: Joel Johansson

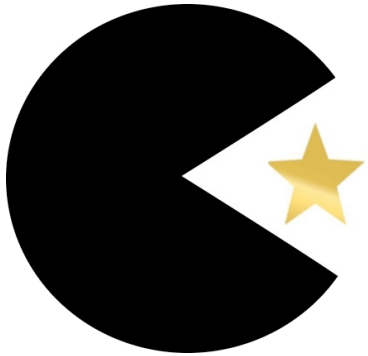


ZTF is an optical telescope with a 47 sq. deg. field of view

ZTF surveys the entire northern sky every 2 nights, in g+r, as part of a public survey

The ZTF TDE program

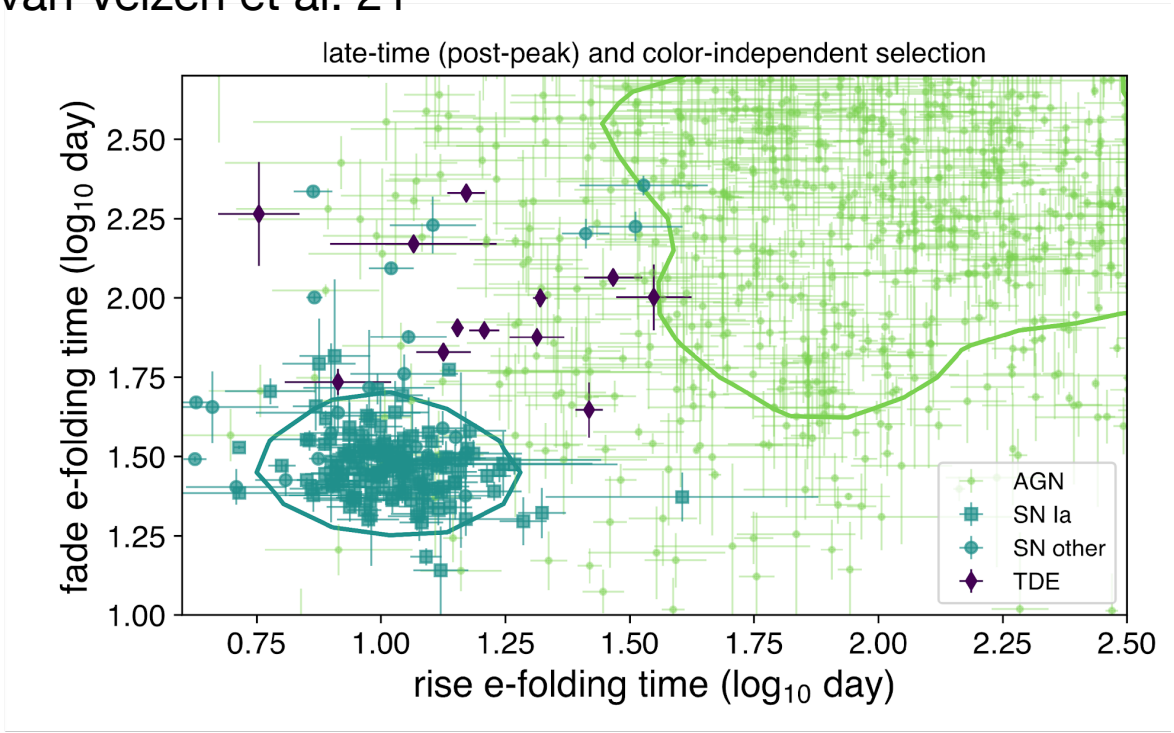


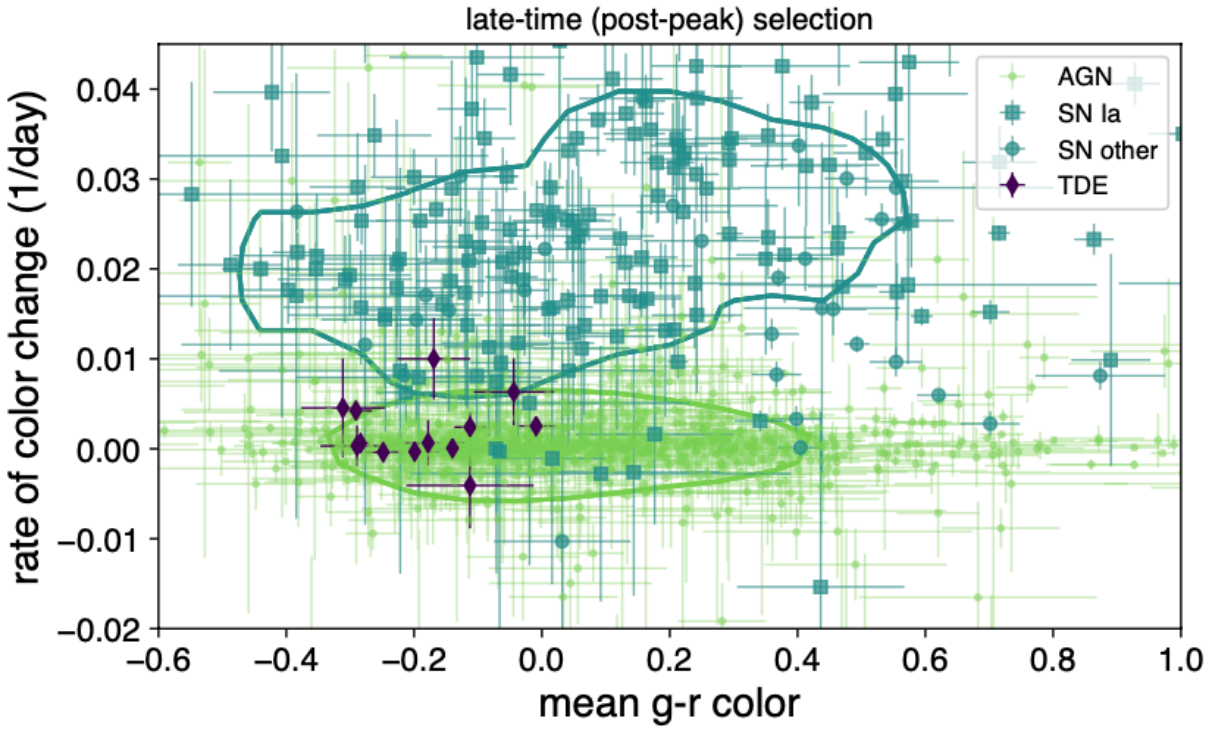


“No TDE left behind”
- since 2018

The ZTF TDE program

van Velzen et al. 21

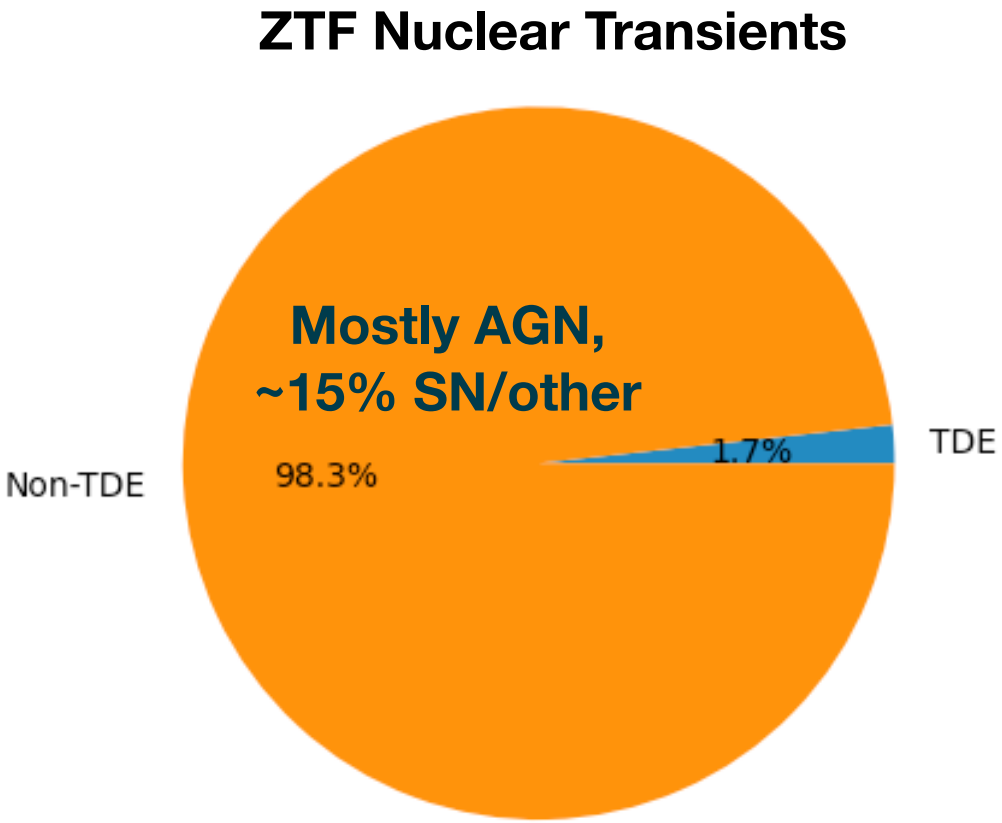




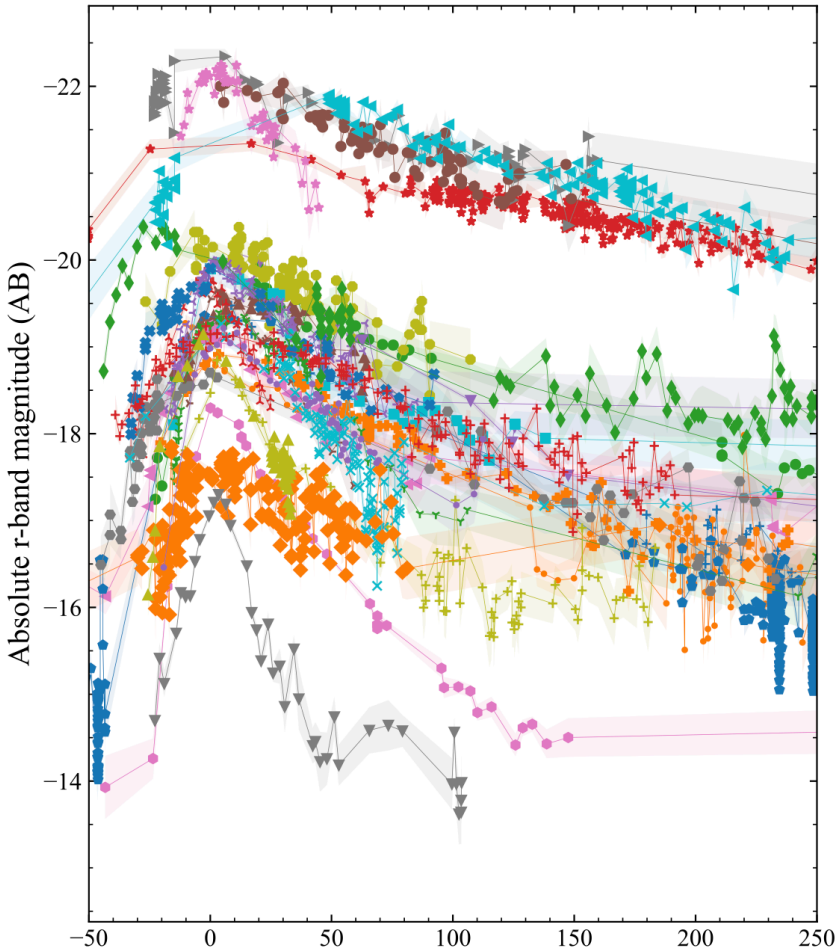
ZTF has a systematic program to identify and classify nuclear transients

~11000 sources passed latest iteration of filter (Reusch et al. in prep), after which we do mag-based census

The ZTF Nuclear Sample



Hammerstein et al. 23



ZTF has high-cadence two-band coverage of ~80 TDEs

Also have an extensive sample of TDE imposters, among ~5000 classified sources

TDEs with Rubin: from dozens to thousands?

With multi-colour light curves and exquisite depth with TDE-appropriate cadence, Rubin should soon start detecting thousands of TDEs

However, mass spectroscopic classification programs will not be feasible

Credit: Gianluca Lombardi

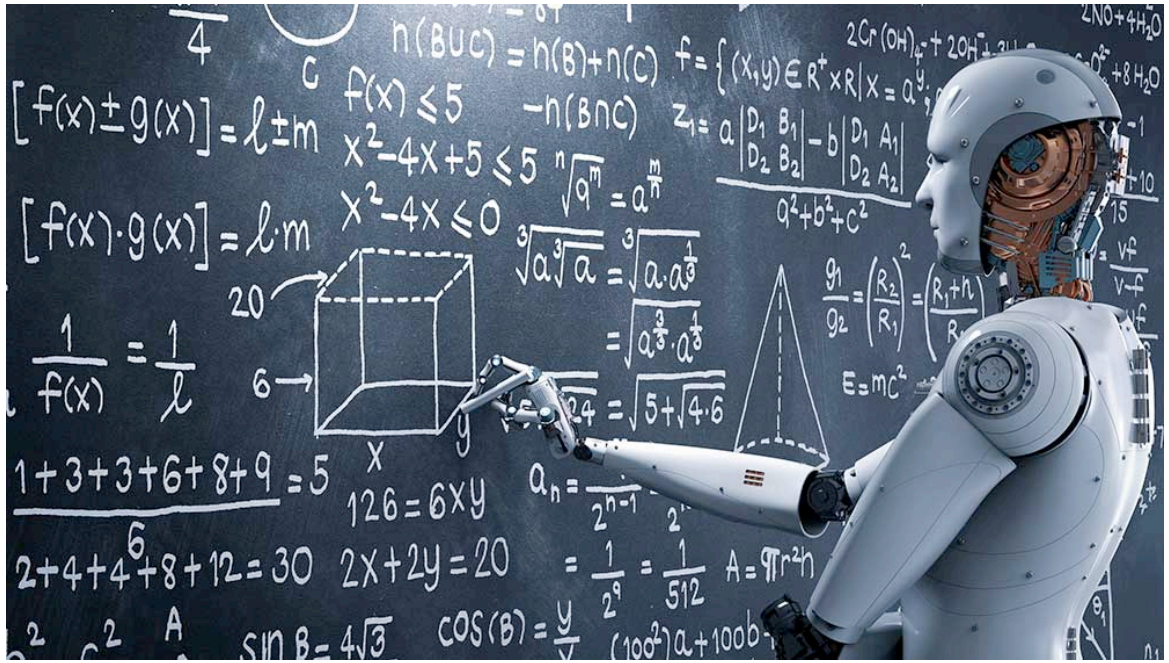
The solution?

NYT, March 28, 2023

Automating astronomy

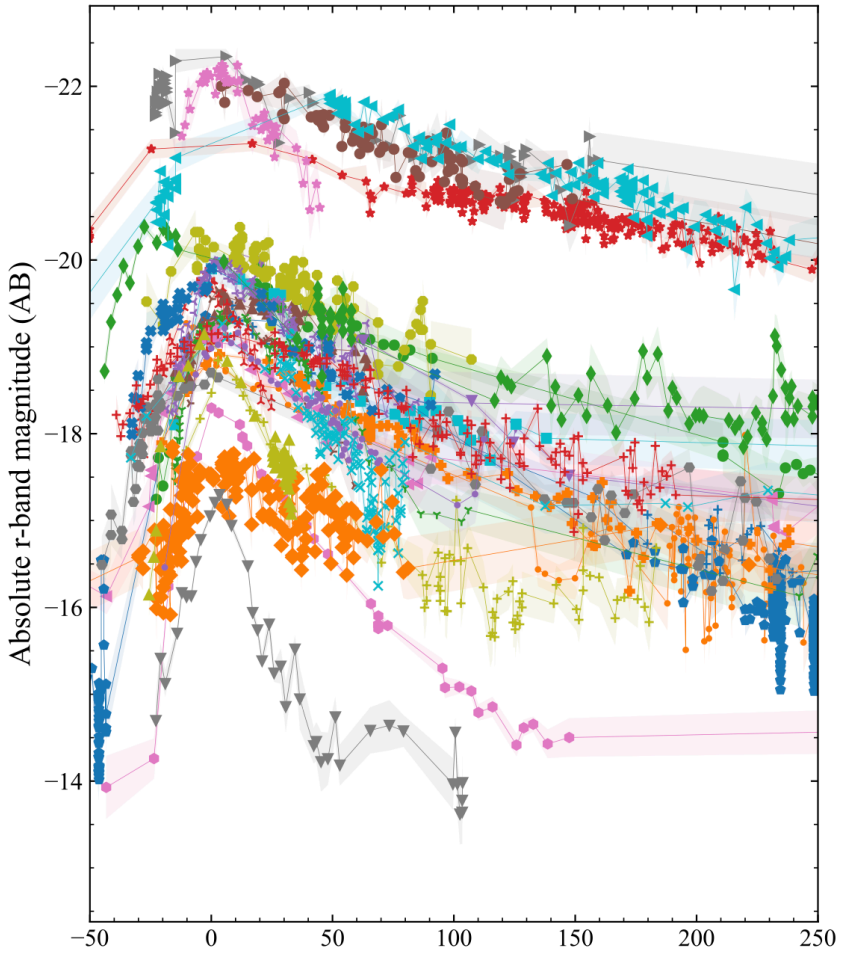
Replacing astonomers with AI: a three step guide

1. Make problem “machine-readable”
2. Develop ML model tailored to problem
3. Train, test, deploy



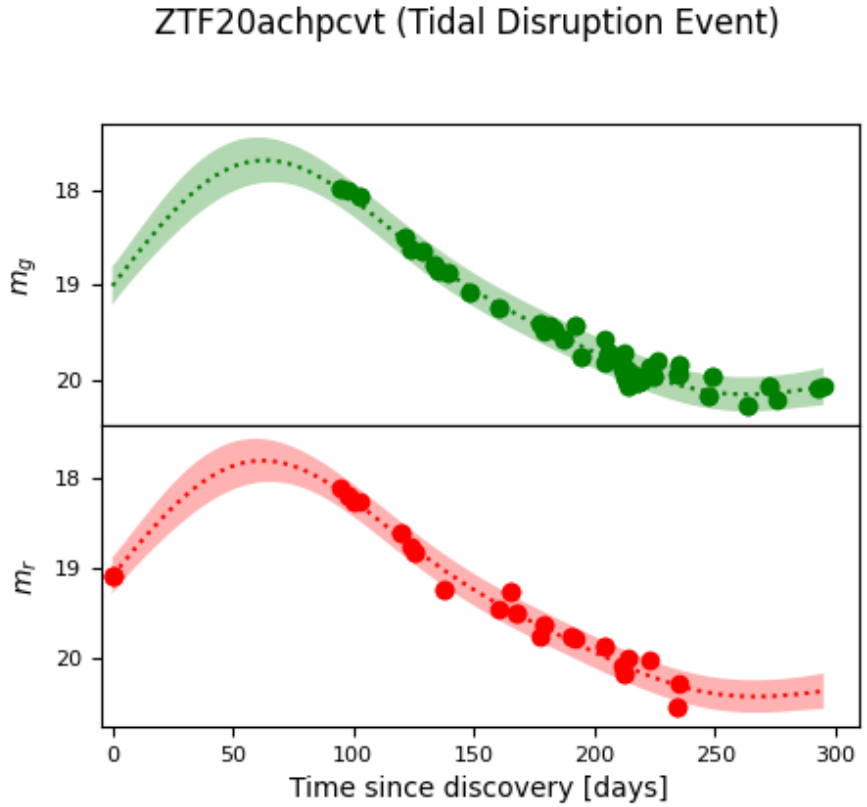
Step 1: Make the ZTF data “machine-readable”

Hammerstein et al. 23



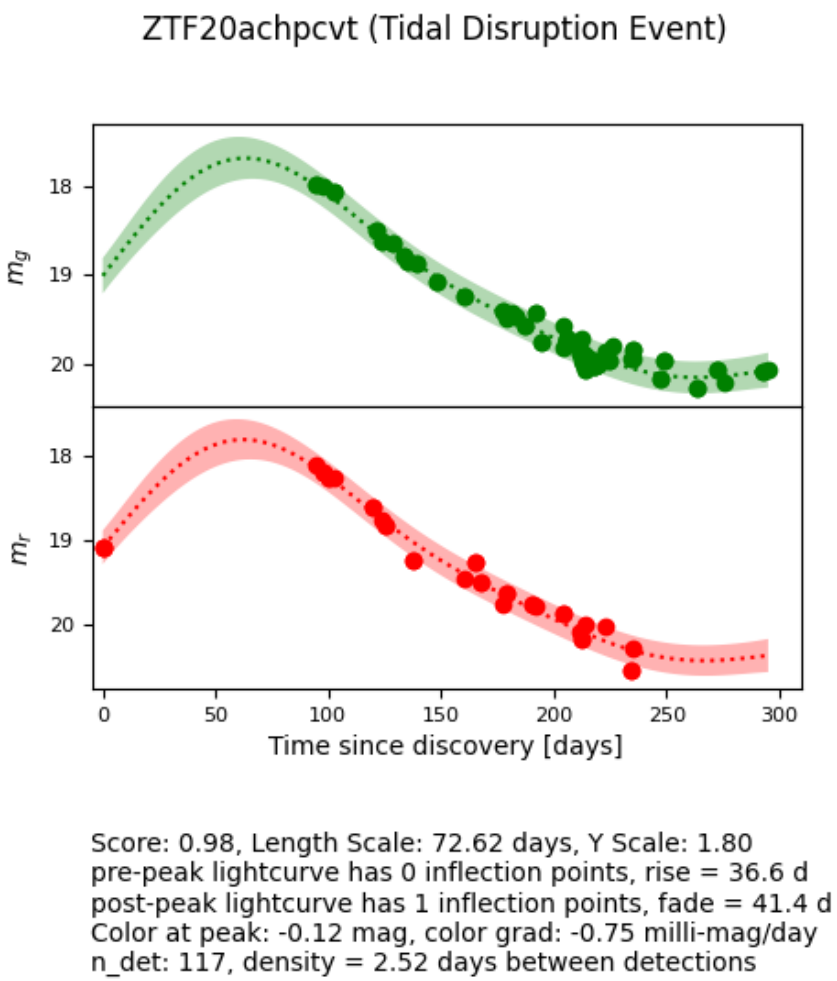
~5000 classified nuclear sources from ZTF

Use Gaussian Process with linear colour evolution



Interpolated light curves for thousands of sources

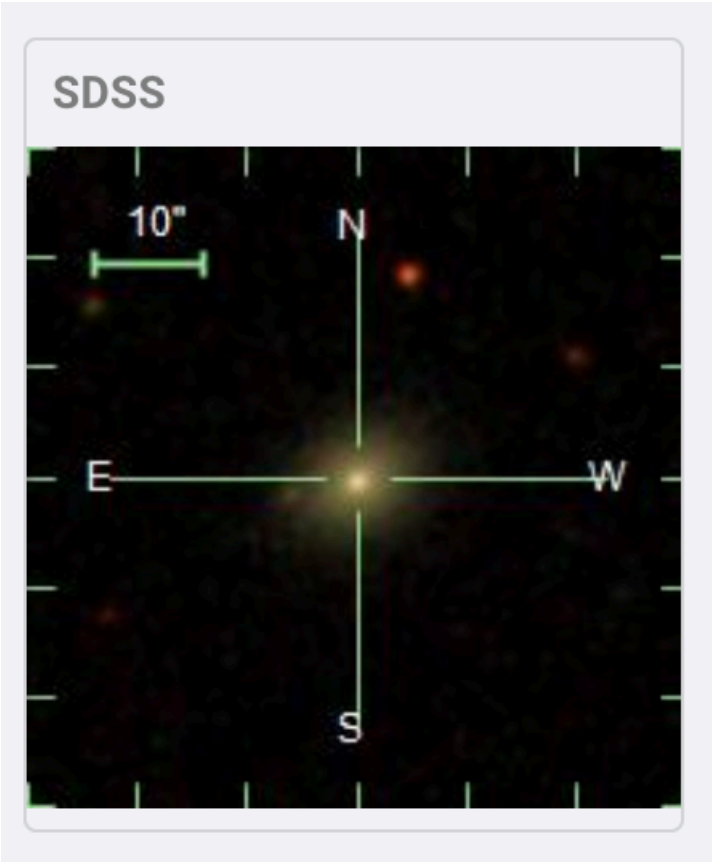
Step 1: Make the ZTF data “machine-readable”



- With the gaussian process fits:
  - Extract peak
  - Extract rise time/fade time (peak-0.5 mag)
  - Extract peak colour, and colour gradient
  - Count inflection points (pre, post) - AGN are bumpy
  - Score - How well the model can capture data
  - Y scale/amplitude increase

For fun:
SNcosmo fit as a proxy for Ia-ness

Step 1: Make the ZTF data “machine-readable”



- Data from crossmatch:
  - WISE host colours
  - PS1 host colours
  - Milliquas
- Info from the ZTF alerts themselves:
  - star galaxy score,
  - Distance to host
  - fraction of positive detections

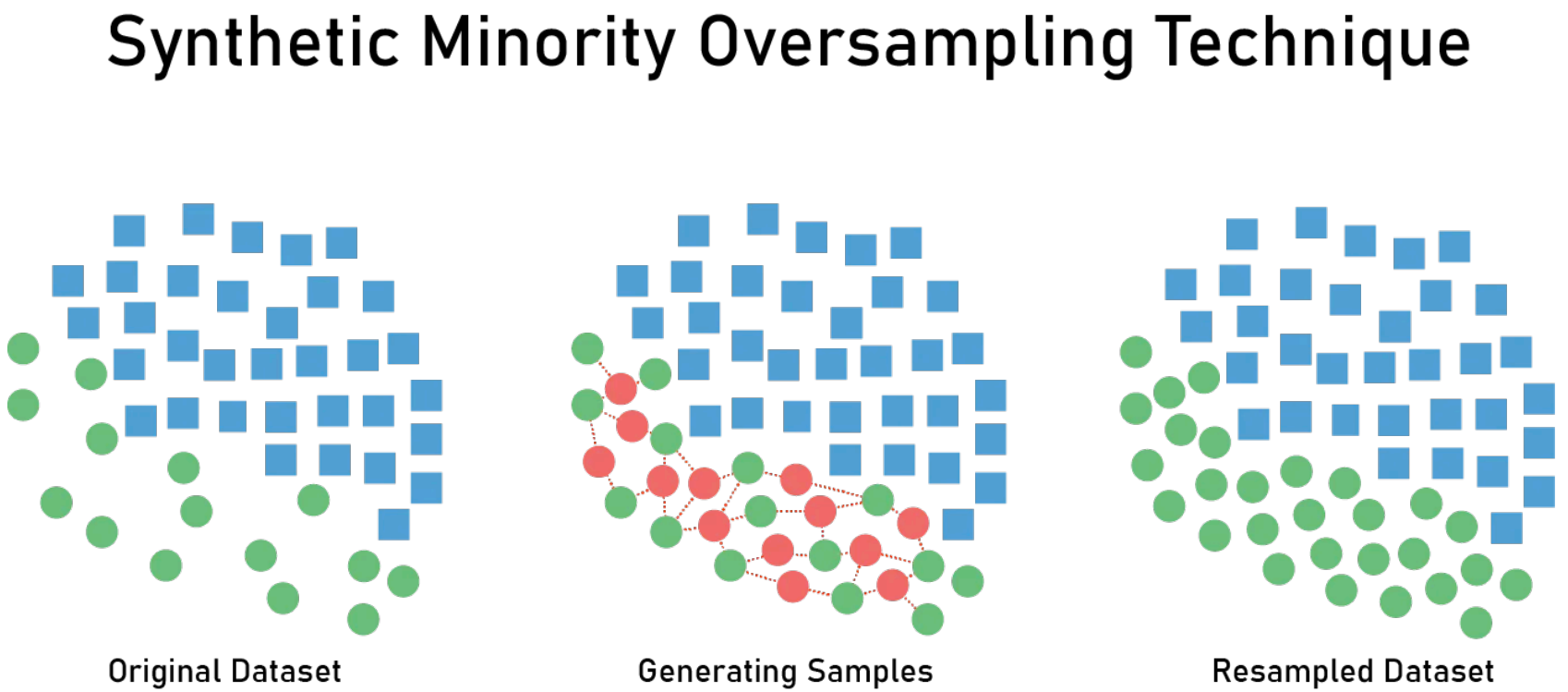
Ultimately have ~50 TDEs with well-sampled rise/fade/colour and all relevant crossmatches, alongside ~2500 other nuclear transients

Step 2: Tailor ML solution to specific problem

The problem:

f(x) = “Not TDE” **-> Will be ~98.5% accurate**

The solution:

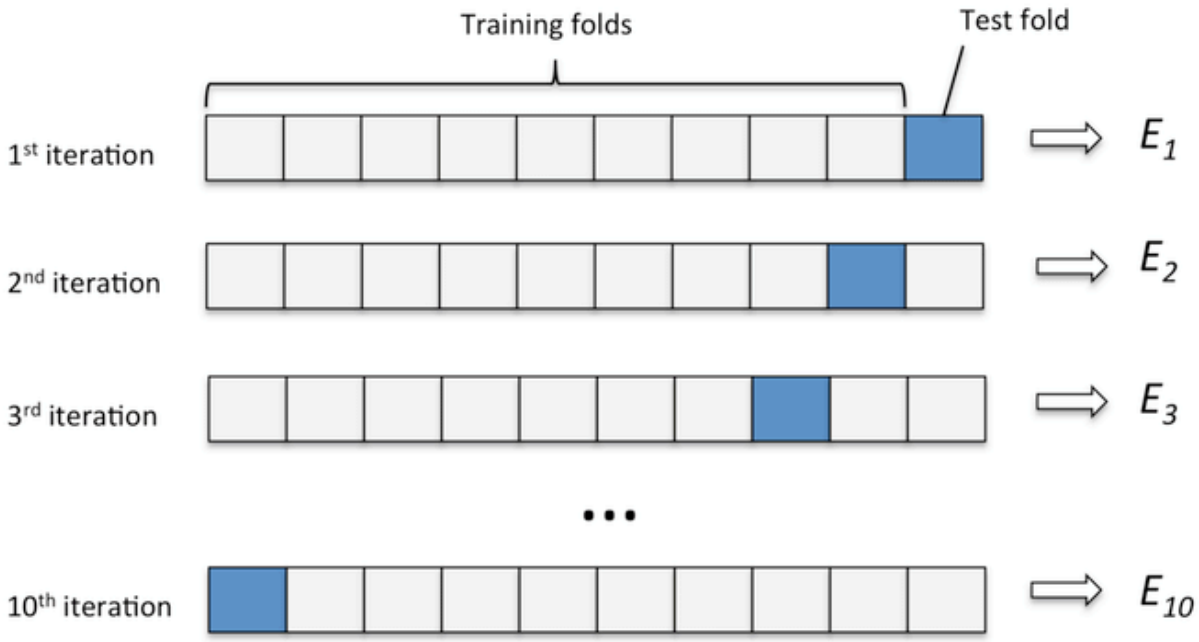


Credit: Y Charfaoui

“tdescore”

Step 3: An ML architecture for photometric classifier

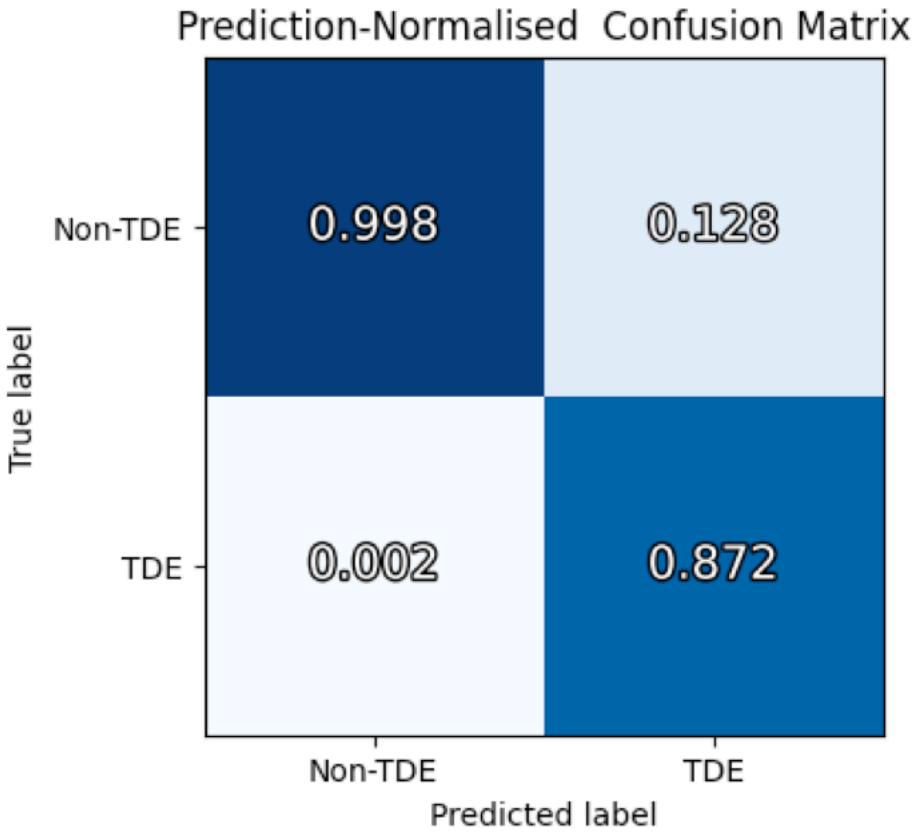
- Our problem is not particularly complicated
  - Use a simple “XGBoost” classifier
- Only a handful of TDEs in total, and no reliable simulation method for lightcurves
  - Use k-fold cross validation to measure performance (“leave one out method”)
- Small dataset (5000 sources)
  - Do not heavily optimise hyper parameters to avoid overtraining

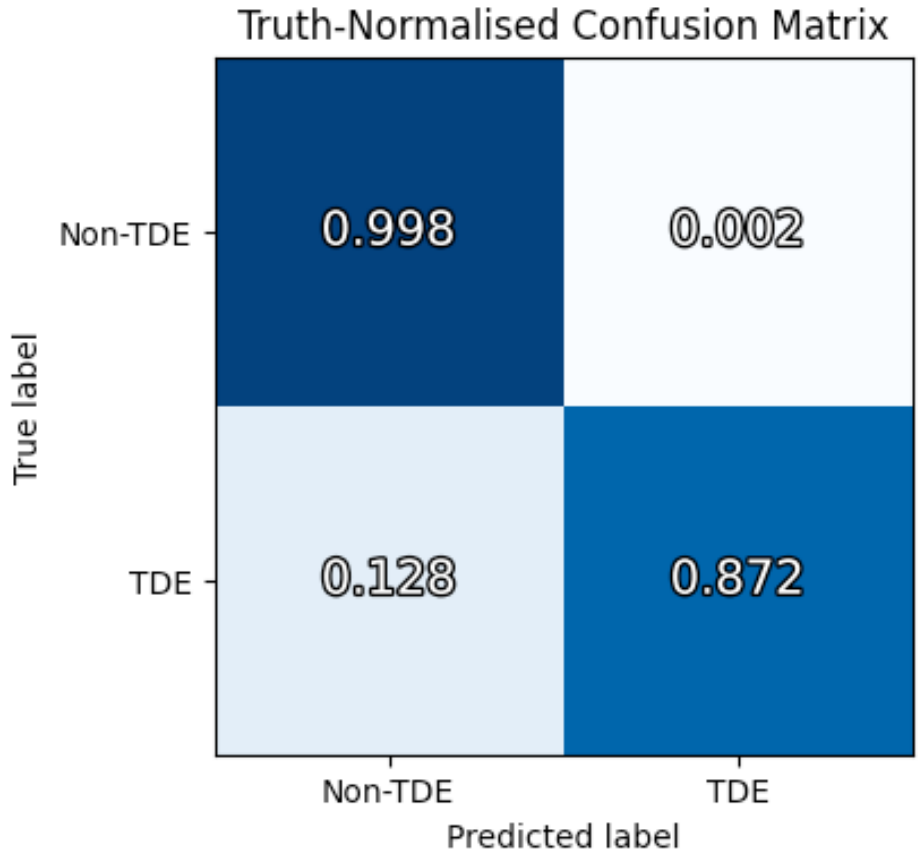


See https://github.com/robertdstein/tdescore

Credit: K Rosaen

The result: an accurate photometric classifier





Rejects non-TDEs with 99.8% efficiency

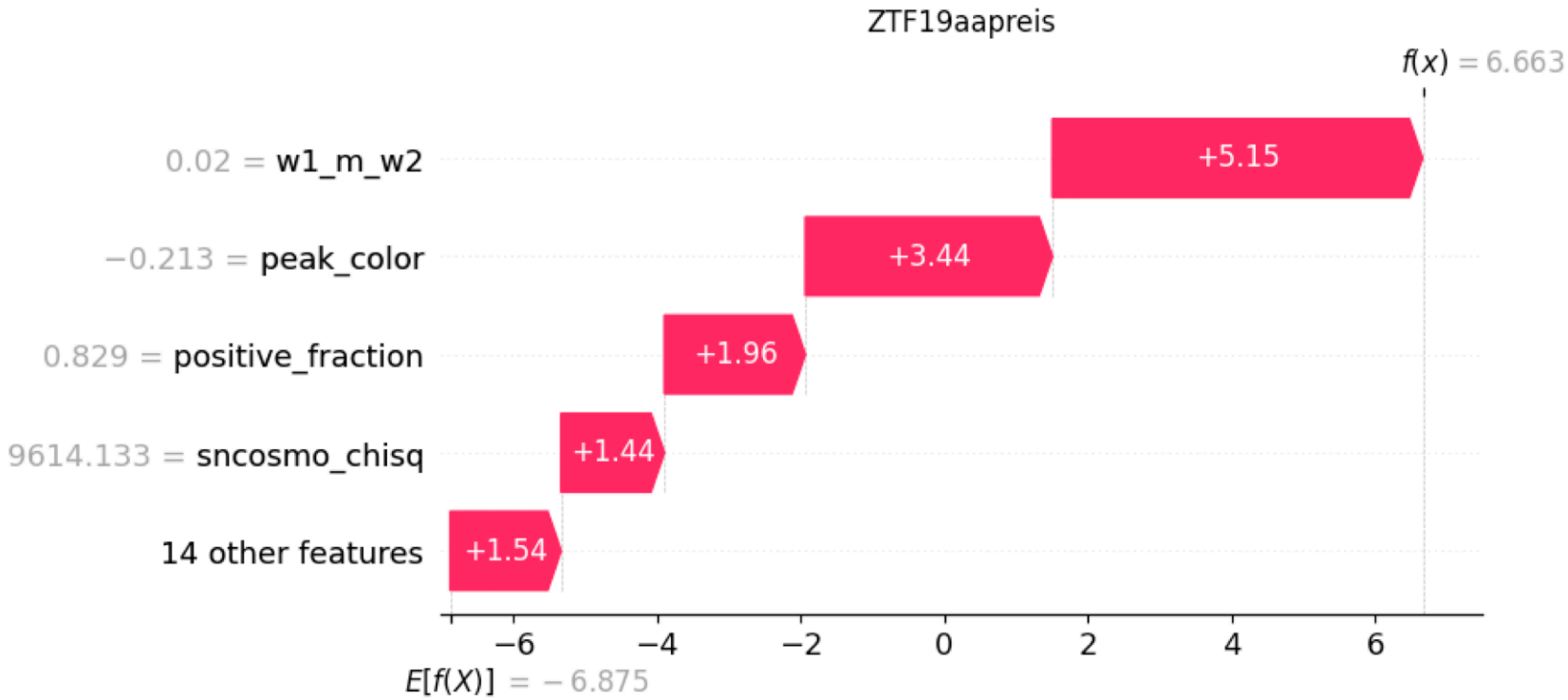
Yields very pure (>80%) sample of TDEs

Explainable AI

Which you can argue with!

Real TDE gets high score

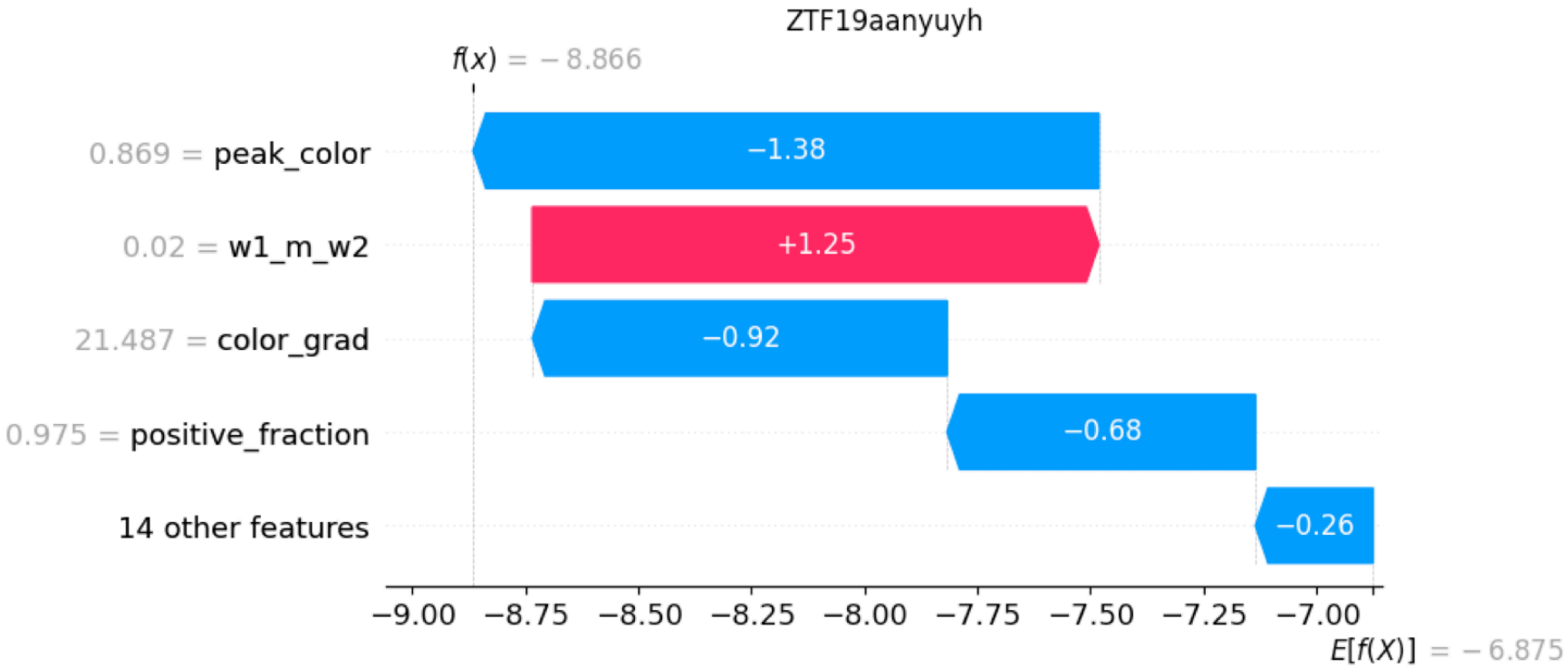
(Due to e.g non-AGN WISE colour, peak colour…)



Which you can argue with!

Real SNIa gets low score

(due to colour at peak, cooling …, and despite non-AGN WISE colour)



Summary

Summary

- ZTF has a systematic program of TDEs with high-cadence light curves. Unprecedented sample of TDEs, but also of TDE imposters!
- Using XGBoost random forest and the ZTF nuclear sample, I extracted parameters for ~5000 sources and trained a photometric classifier, **tdescore**
- **tdescore** was able to reject non-TDEs with 99.8% accuracy, leaving pure sample of mostly (>80%) TDEs
- Explainable AI method helps us to understand/ evaluate classifier reasoning.
- Expect to see the paper on arXiv in July!

Credit: Dall-e