



*XMM2Athena Workshop: XMM Newton survey legacy for Athena and beyond
IRAP, Toulouse - France
28 February 2024*

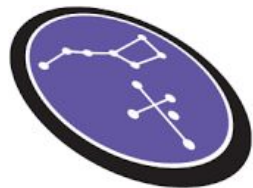
Emille Ishida, Anais Möller and Julien Peloton
on behalf of the Fink Team

MMA and discovery with Rubin

- State of the art data sets in optical astronomy
- The Rubin broker ecosystem
- The role of Machine Learning
- The human factor

How did we get here?

Big data happened



SDSS

2000 - now

120 TB

Primary mirror: 2.5 m



ZTF

2018 - now

1.4 TB/night

Primary mirror: 1.2 m



**VERA C. RUBIN
OBSERVATORY**

15 TB/night

From 2025

Primary mirror: 8.4 m

The Vera Rubin Observatory Large Survey of Space and Time (LSST)

In a nutshell:

- telescope: 8.4m primary mirror
- world's largest CCD camera:
3.2 Gpixels

In numbers:

- 10-year survey, starting 2024+
- 1,000 images/night = 15TB/night
- 10 million transient candidates per night
 - Publicly available...
 - ... but huge!



To keep in mind ...

- 1) Very big data
- 2) Survey mode observation strategy

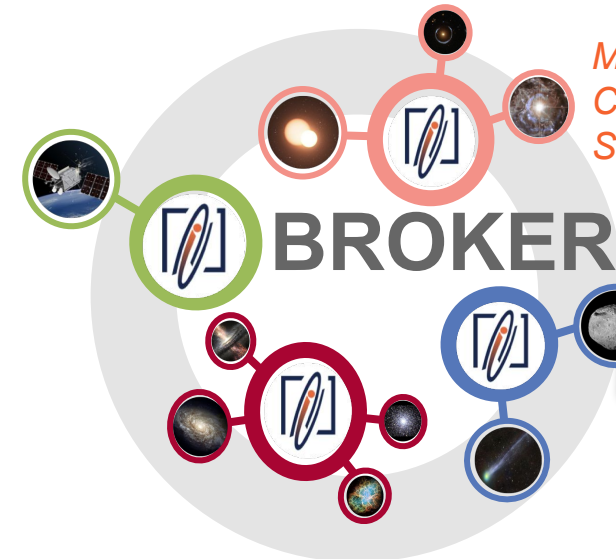
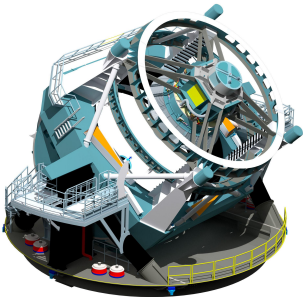
How to distribute this to people who will do science?

From detection to science

The data path



every ~30 seconds down to mag ~24



*Machine learning
Catalog association
Streams join*



10 million alerts per night...



We would like the interesting ones ...



Rubin broker landscape



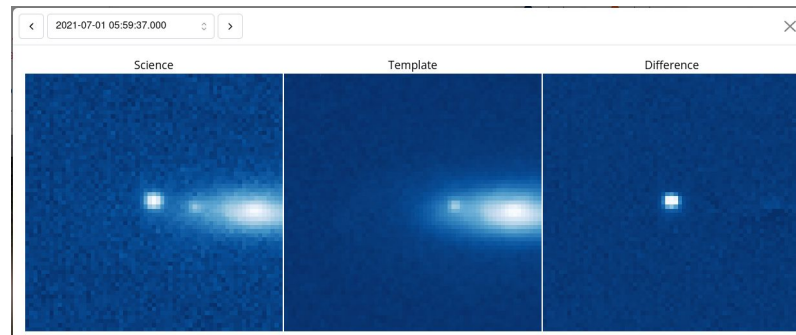
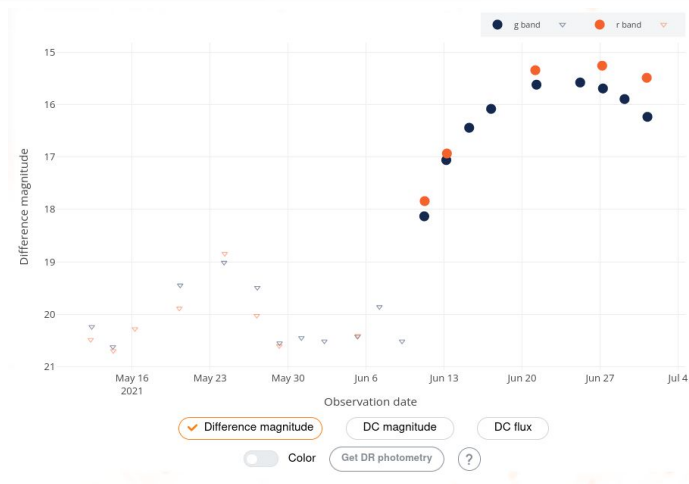
(What is an alert?)



Alerts based on Difference Image Analysis

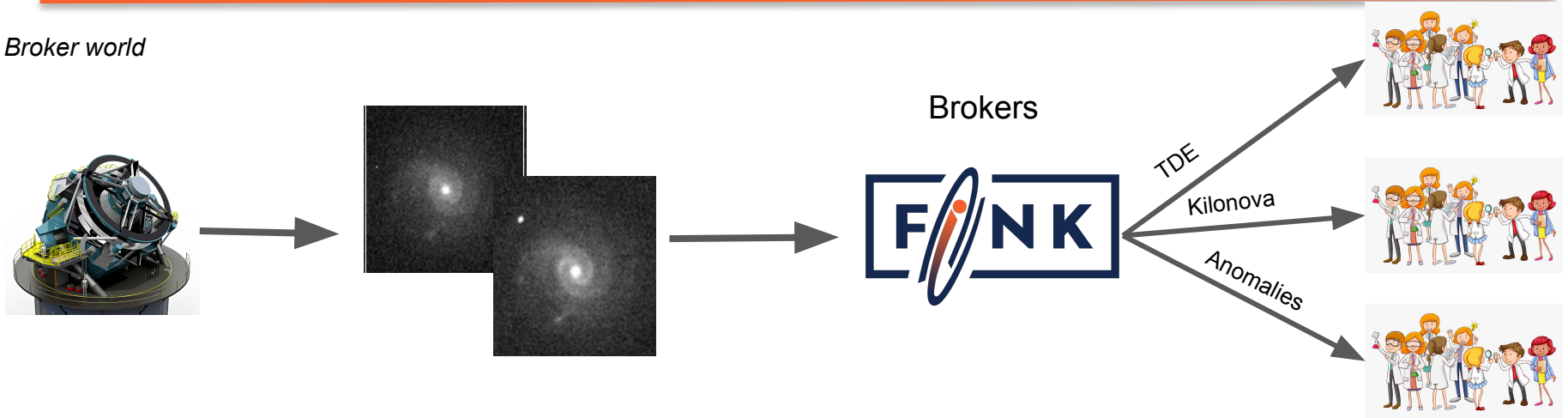
Each alert contains

- Information about the new detection (magnitude, position, ...)
- Neighbours information (xmatches, etc)
- Historical information if the object has been seen previously
- Small images around the detection (60x60 pixels)



How to classify alerts?

Broker world

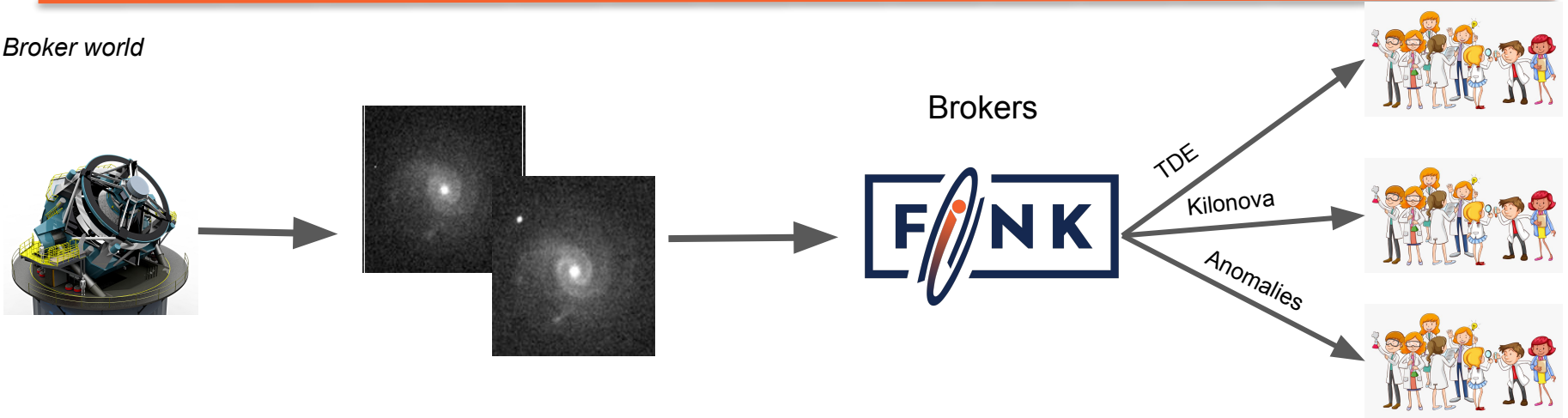


Domain specialist world (this is you)

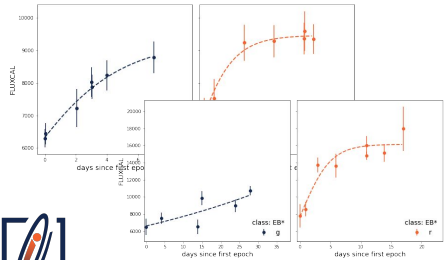


How to classify alerts?

Broker world



Domain specialist world (this is you)



You can access this via the Fink Science portal or the API

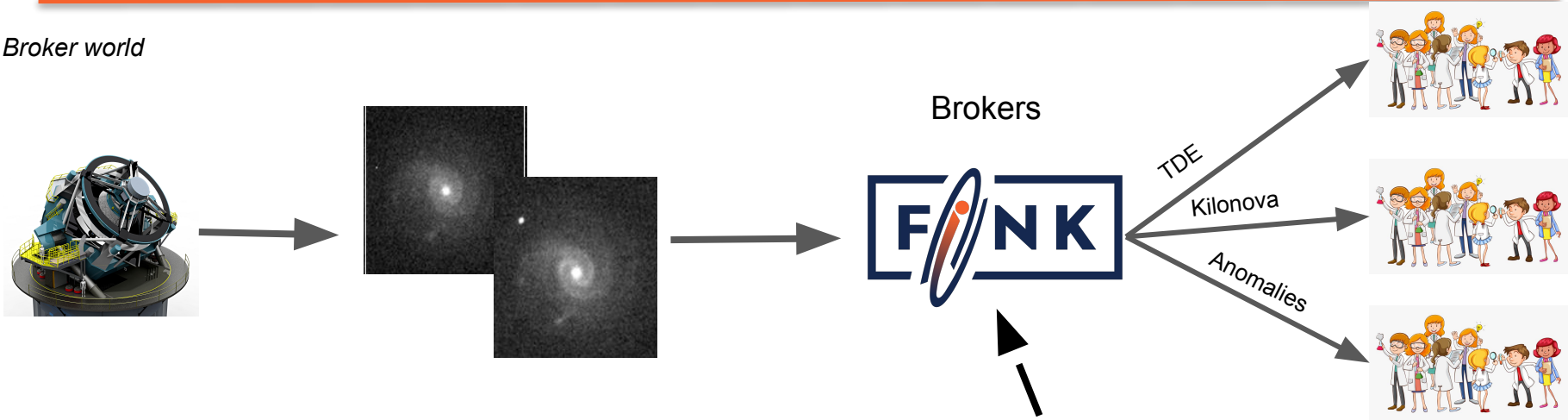
All alerts data is public!

<https://fink-portal.org/>

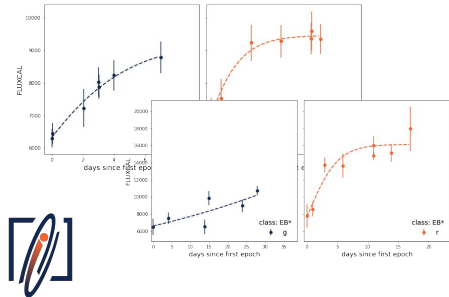


How to classify alerts?

Broker world



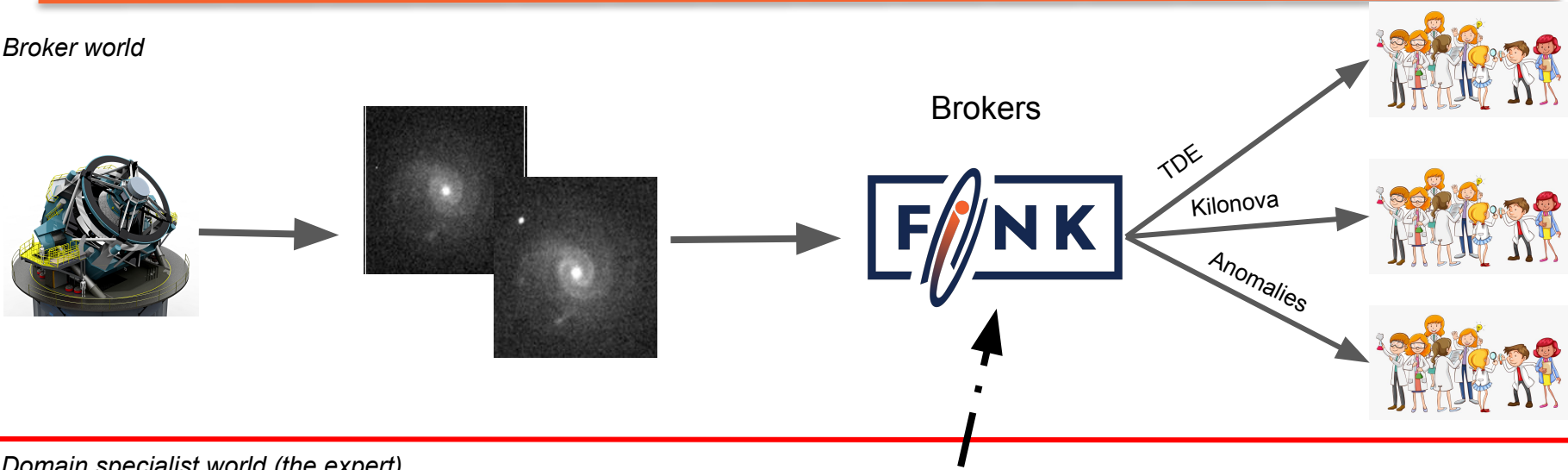
Domain specialist world (the expert)



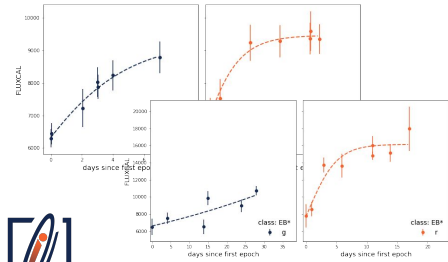
Filter
Catalog or stream xmatch
Tailored science module

How to classify alerts?

Broker world



Domain specialist world (the expert)



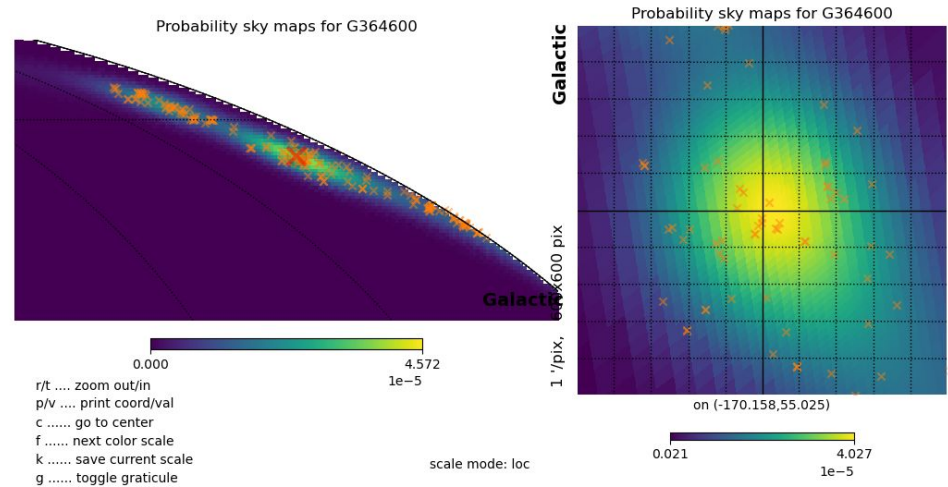
- Magnitude
- Position, or any column really
- X-match with
 - existing catalogs
 - other streams
 - probability maps

Xmatch with GW streams

O4 is coming – Fink has already some tool to play with GW sky maps

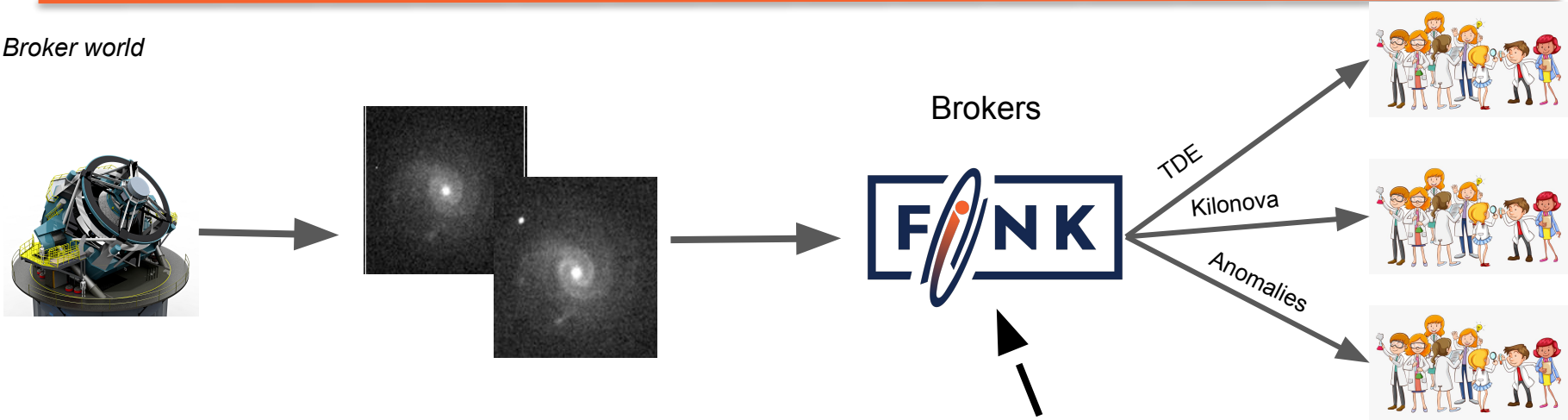
<https://fink-portal.org/api> → Gravitational Waves → tutorial!

```
# Query Fink
data = gzip.open(fn, 'rb').read()
r = requests.post(
    'https://fink-portal.org/api/v1/bayestar',
    json={
        'bayestar': str(data),
        'credible_level': credible_level,
        'output-format': 'json'
    }
)
```

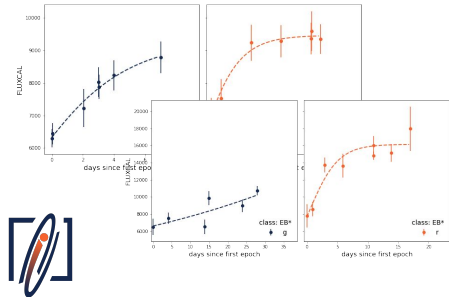


How to classify alerts?

Broker world



Domain specialist world (the expert)

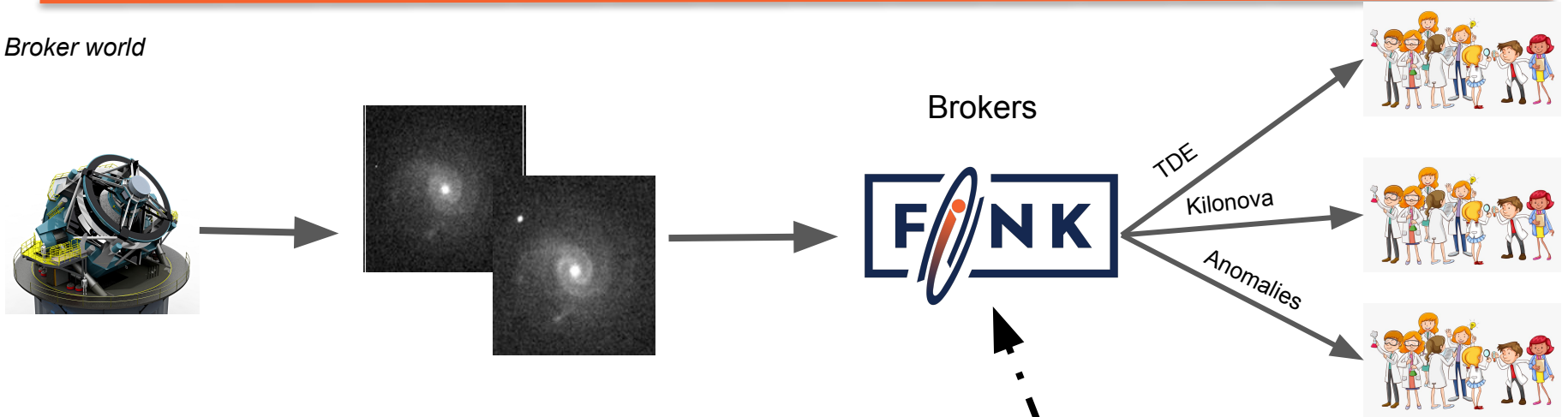


Taylorred science module

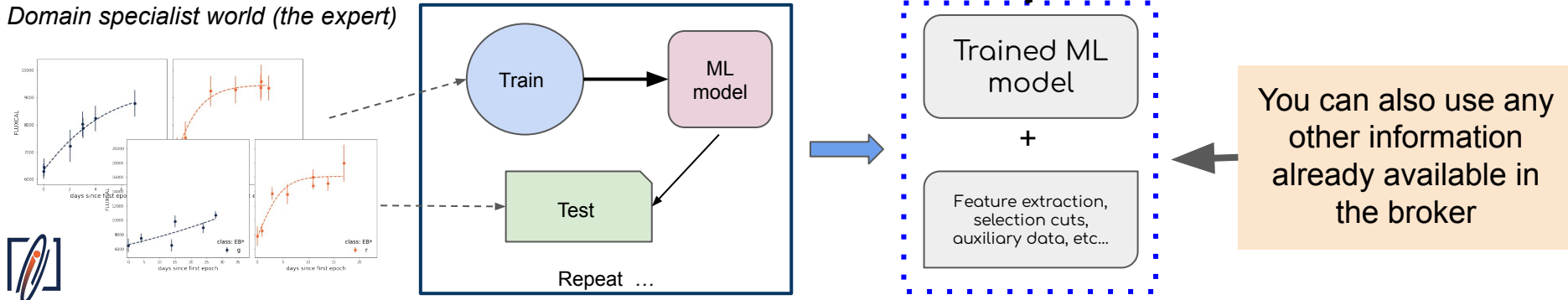
$f(\text{alerts}; ++)$ => *class scores*
Boolean

How to classify alerts?

Broker world



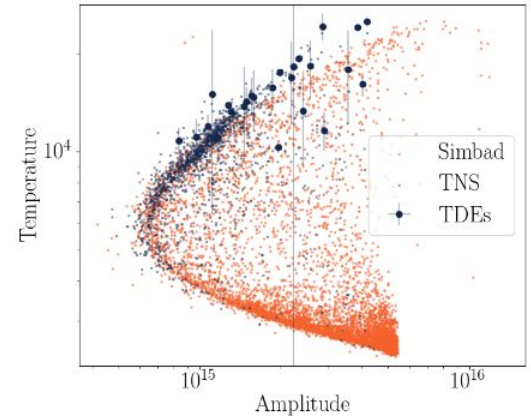
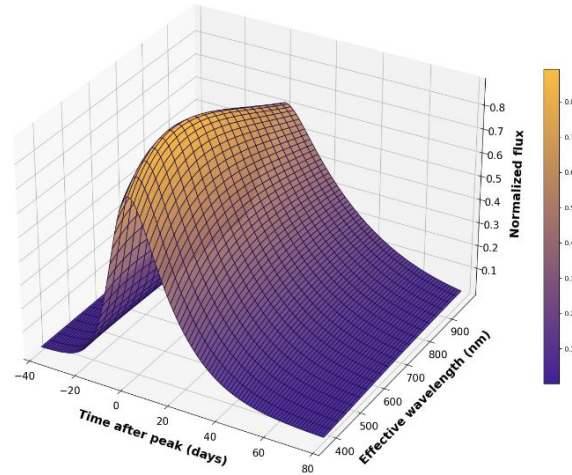
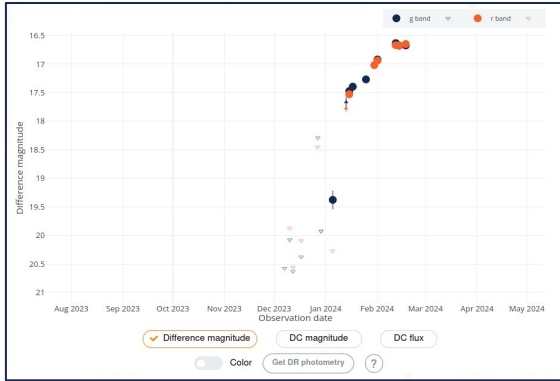
Domain specialist world (the expert)



Example: TDE



- Focused on rising TDE examples
- Remove things we already know (xmatch with whatever possible)
- Multi-wavelength feature extraction with Rainbow
- Anomaly detection based ML-model



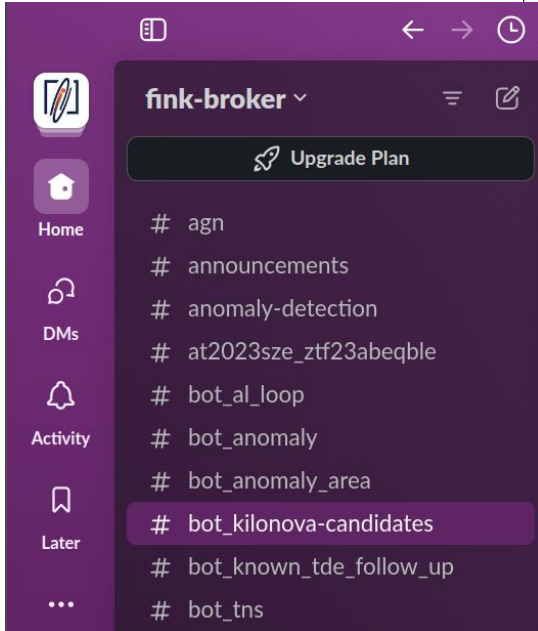
Receiving your candidates

- Kafka stream -- for on-the-fly notifications



Receiving your candidates

- Bots



Cross-match-based kilonova bot APP 9:16 AM

Fink Science Portal: [ZTF24aaemydm](#)

SkyPortal: [ZTF24aaemydm](#)

Time:

- 2024-02-12 08:02:33.003 UTC
- Time since first detection: 1.9 hours

Presumed host galaxy:

- HyperLEDA Name: NGC3052
- 2MASS XSC Name: 09542791-1838202
- Luminosity distance: (57.25 ± nan) Mpc
- RA/Dec: [148.6163330](#) -18.6389450
- log₁₀(Stellar mass/Ms): 10.81

RA/Dec:

- [hours, deg]: 9 54 28.51 -18 38 10.9
- [deg, deg]: [148.6188113](#) -18.6363575

Cross-match:

- Alert-host distance: 3.43 kpc
- Absolute magnitude: -15.54

Galactic latitude:

- [deg]: 27.2874401

Measurement (band r):

- Apparent magnitude: 18.25 ± 0.07

TNS: [link](#)



Receiving your candidates

- Added value + download service

Select data source
Source: ZTF

Filter alerts
Dates: 2023-06-01 - 2024-01-31
Classe(s): ['SN candidate', '(TNS) SN', '(SIMBAD) SN']
Conditions: nalerthist>5;

Select content
Content: Full packet

Submit
Trigger your job!

 Description



nalerthist>5; magpsf < 22

 Log in



Alert content

Choose the content you want to retrieve

Lightcurve (~1.4 KB/alert) Cutouts (~41 KB/alert) Full packet (~55 KB/alert)

 Submit job

Fink Data Transfer

Data Source

Choose the type of alerts you want to retrieve

ZTF ELASTICC (v1) ELASTICC (v2.0) ELASTICC (v2.1)

Filters

Date Range *

Pick up start and stop dates (included).

June 1, 2023 – January 31, 2024

Alert class

Select all classes you like! Default is all classes.

(Fink) Supernova candidates x (TNS) SN x (SIMBAD) SN x

Extra conditions

One condition per line (SQL syntax), ending with semi-colon. See [here](#) for fields description and [here](#) for examples.

nalerthist>5; magpsf < 22

Alert content

Choose the content you want to retrieve

Lightcurve (~1.4 KB/alert) Cutouts (~41 KB/alert) Full packet (~55 KB/alert)

 Submit job



The beauty of an observational science

“... telescopes that merely achieve their stated science goals have probably failed to capture the most important scientific discoveries available to them.”

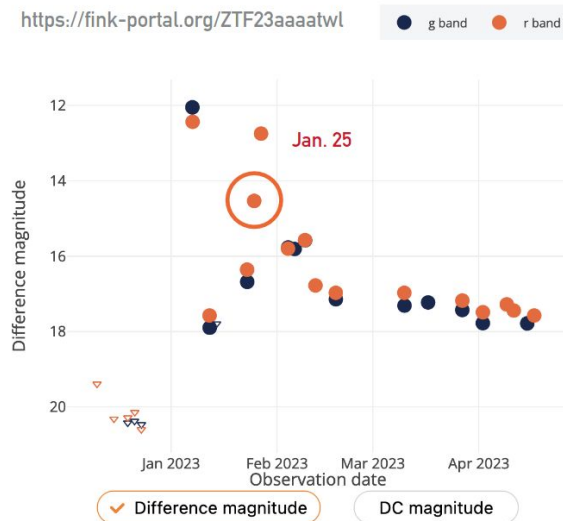
Norris, R. (2017). Discovering the Unexpected in Astronomical Survey Data. Publications of the Astronomical Society of Australia, 34, E007. doi:10.1017/pasa.2016.63

Fink Anomaly Detection module

- Started with extragalactic experts, now Fink-wide engagement
- Random forest model
- Quick reaction from the follow-up community:
 - 9.2m SALT (South Africa)
 - 0.6m and 2.5m KGO (Russia)
 - 0.25m FRAM-ORM (Spain)
 - 0.2M FOSC-ES32 (Italy)

FIRST NIGHT – FIRST ANOMALY

<https://fink-portal.org/ZTF23aaaatwl>



▶ ASASSN-23ac/PNV
J06245297+0208207

▶ Simbad: WD candidate

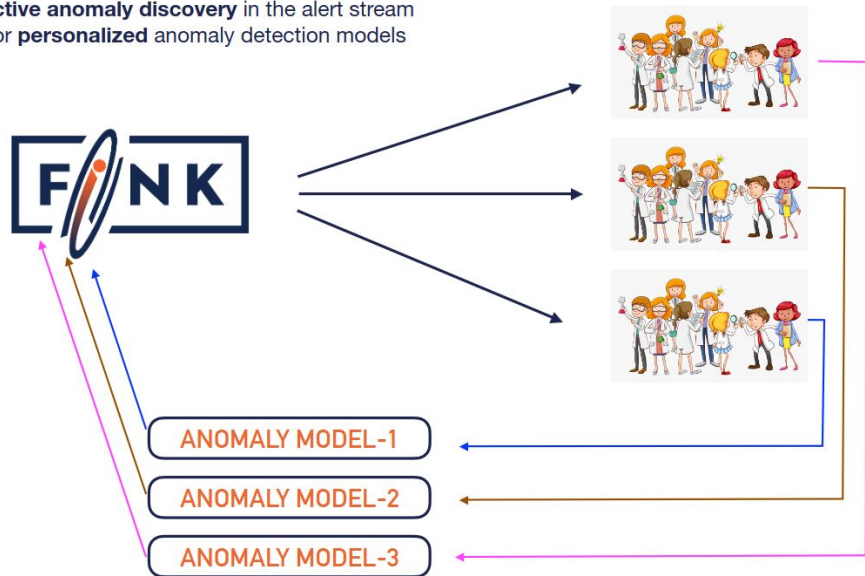


REPORTED IN TNS AS AT 2023AWT



Fink Anomaly Detection module

Active anomaly discovery in the alert stream
for personalized anomaly detection models



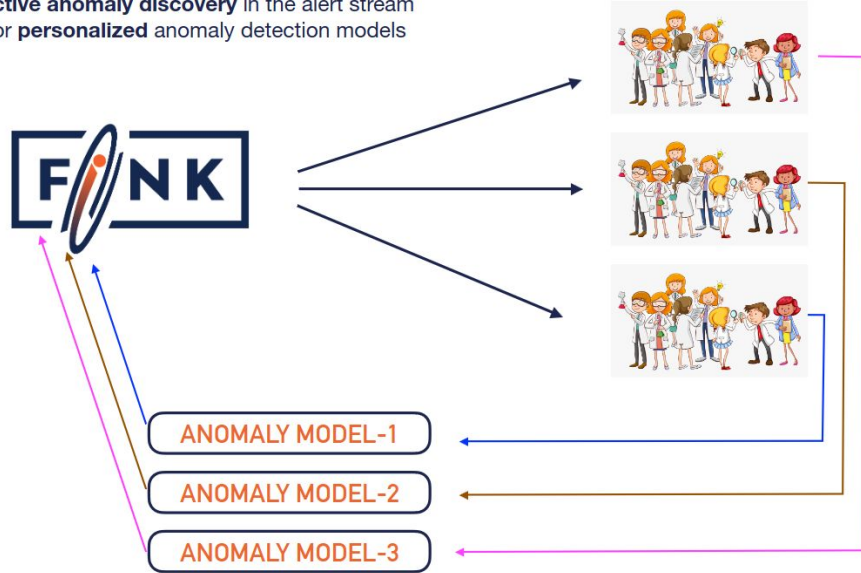
Algorithm from Das, S., et al., 2017, in DEA'17, KDD workshop, [arXiv:cs.LG/1708.09441](https://arxiv.org/abs/cs.LG/1708.09441)

Implementation by the [SNAD team](#), via [coniferest package](#)



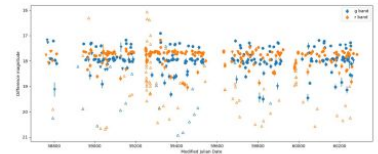
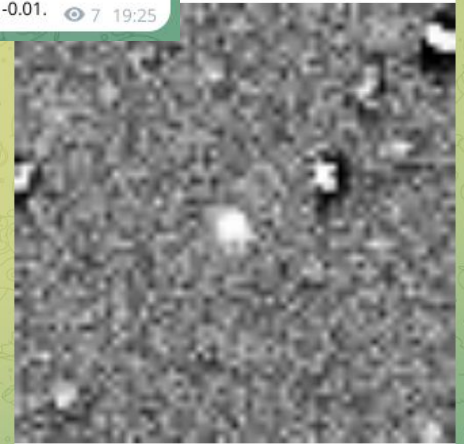
Fink Anomaly Detection module

Active anomaly discovery in the alert stream for personalized anomaly detection models



Fink anomaly bot

Median anomaly score overnight: -0.01. 👁️ 7 19:25



ID: ZTF18aazfbvg
DR OID (<1"): 591105200057212
GAL coordinates: 56.739963, -4.841495
EQU: 298.6919835, 18.6325727
UTC: 2023-12-07 03:12:54.999
Real bogus: 0.68
Anomaly score: -0.04



1

👁️ 7 16:45

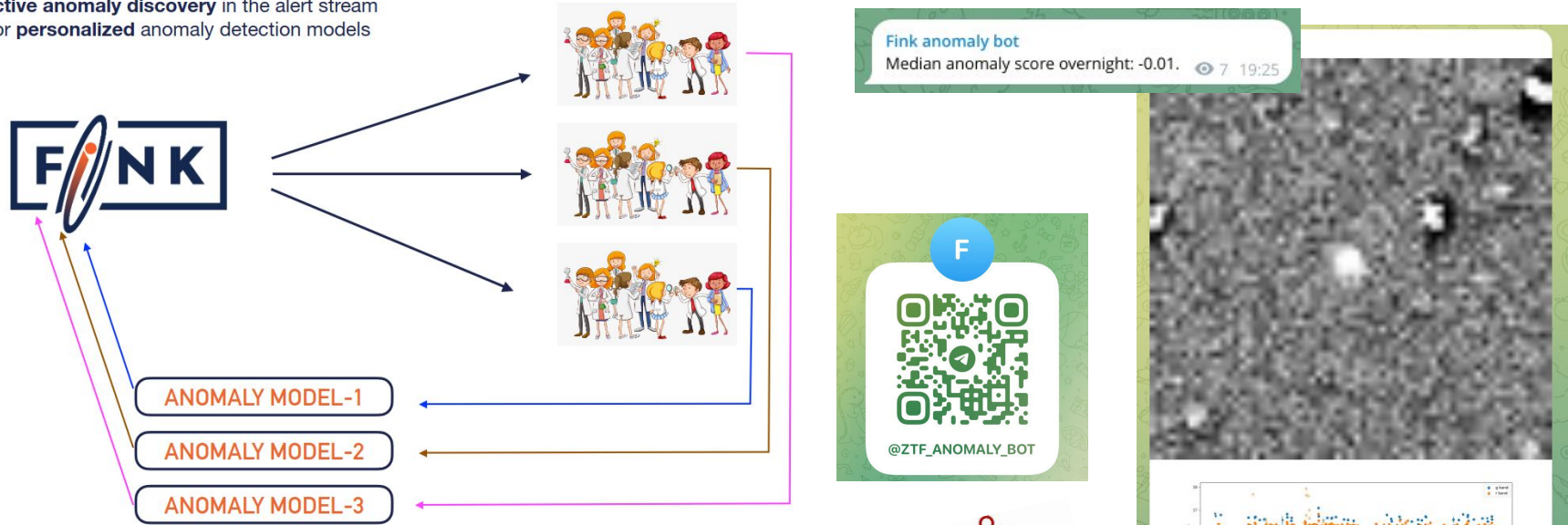
Algorithm from Das, S., et al., 2017, in DEA'17, KDD workshop, [arXiv:cs.LG/1708.09441](https://arxiv.org/abs/cs.LG/1708.09441)

Implementation by the [SNAD team](#), via [coniferest package](#)



Fink Anomaly Detection module

Active anomaly discovery in the alert stream
for personalized anomaly detection models



Algorithm from Das, S., et al., 2017, in DEA'17, KDD workshop, [arXiv:cs.LG/1708.09441](https://arxiv.org/abs/cs.LG/1708.09441)
Implementation by the [SNAD team](#), via [coniferest package](#)



Personalized ML for big data

Fink Collaboration meeting 2024



ozFink meeting 2023



Fink hackathon 2022



Fink collaboration meeting 2022

FinkBR 2024

FINK BRAZIL

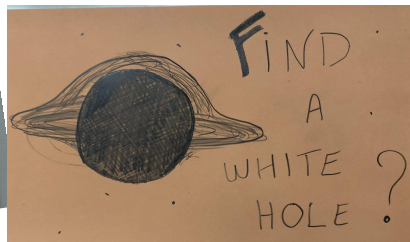
SPEAKERS:
Anais Möller
Swinburne University, Australia
Antonello Palmese
CNR, USA
Claudia Mendes de Oliveira
Universidade de São Paulo, Brazil
Clelio R. Bom
CBPF, Brazil
Charlie Kilpatrick
Northumbria University, UK
Emile E. O. Ishida
CEBRASQ, Brazil
Julien Peloton
CNRS/CEA, France
Mariano Penna-Lima
Universidade de Araraquá, Brazil
Maria Pruzhinskaya
CEBRASQ, Brazil
Martín Mosler
UBA, Argentina and CBPF, Brazil

ENABLING ASTRONOMICAL
TRANSIENT DISCOVERIES
IN THE RUBIN ERA
6 - 10 May 2024
CBPF, Rio de Janeiro

Logos for FINK, CBPF, IAPARU, and others are visible at the bottom of the poster.

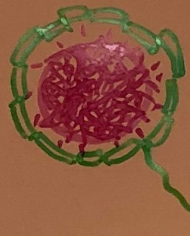
#FinkDreamShots

Build a catalogue
of interstellar asteroids



A classifier
for tidal disruption
events

I WANT TO
FIND A
DYSON
SPHERE



I WANT TO
FIND A LIVE
PISN

~~I wish astronomers~~
~~use REAL units!~~
A switch between
mag. and Lum.

What do you want
to see?

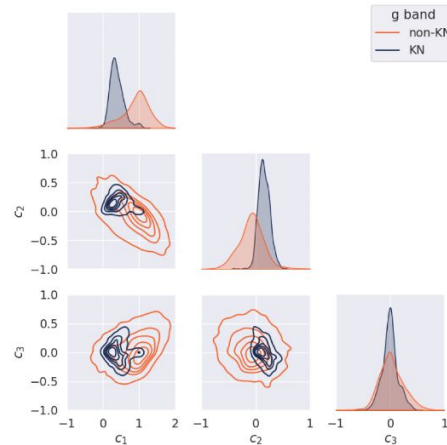
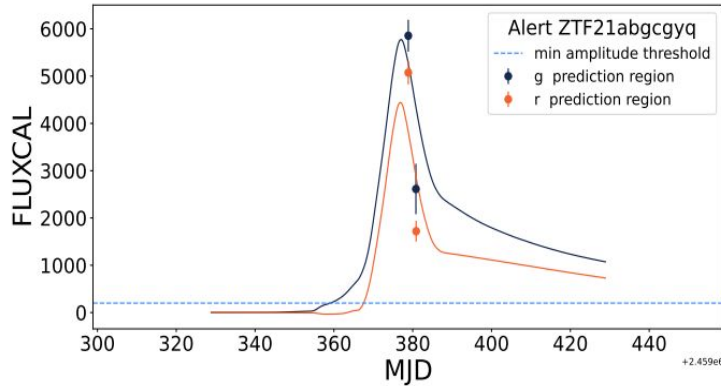




Case study: Kilonova

Problem 1: there are no labels, only 1 confirmed detection- with a GW counterpart

Problem 2: we need to find it fast



<https://fink-portal.org/ZTF21abgcqyq>

Data set:

Simulated ZTF light curves

Feature extraction:

Principal components from perfect sims

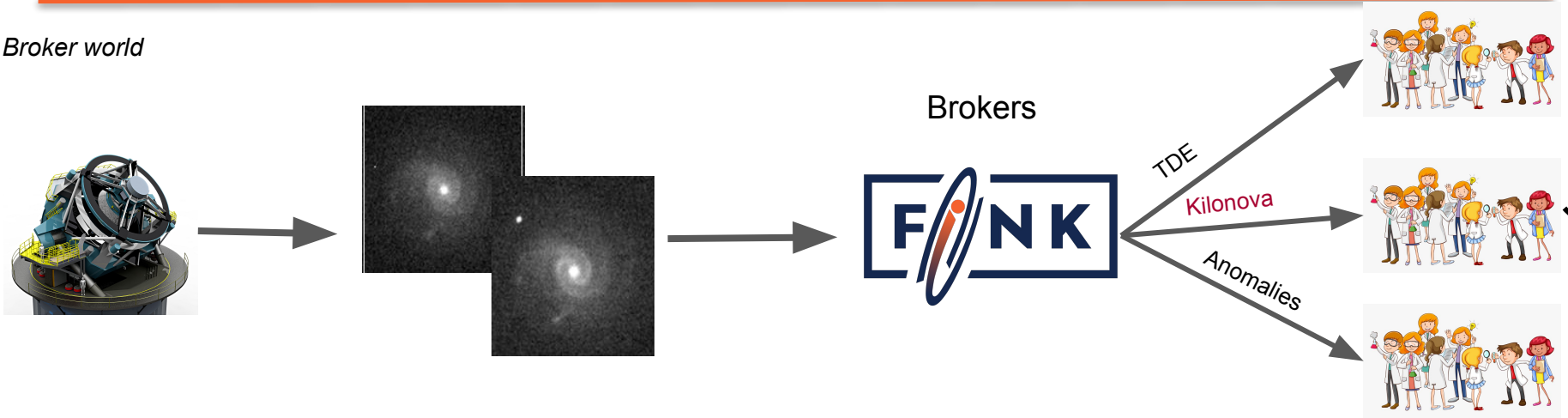
Classifier:

Random Forest



Case study: Kilonova

Broker world



GRANDMA Observations of ZTF/Fink Transients during Summer 2021

Aivazyan et al., 2021, arxiv:astro-ph/2202.09766

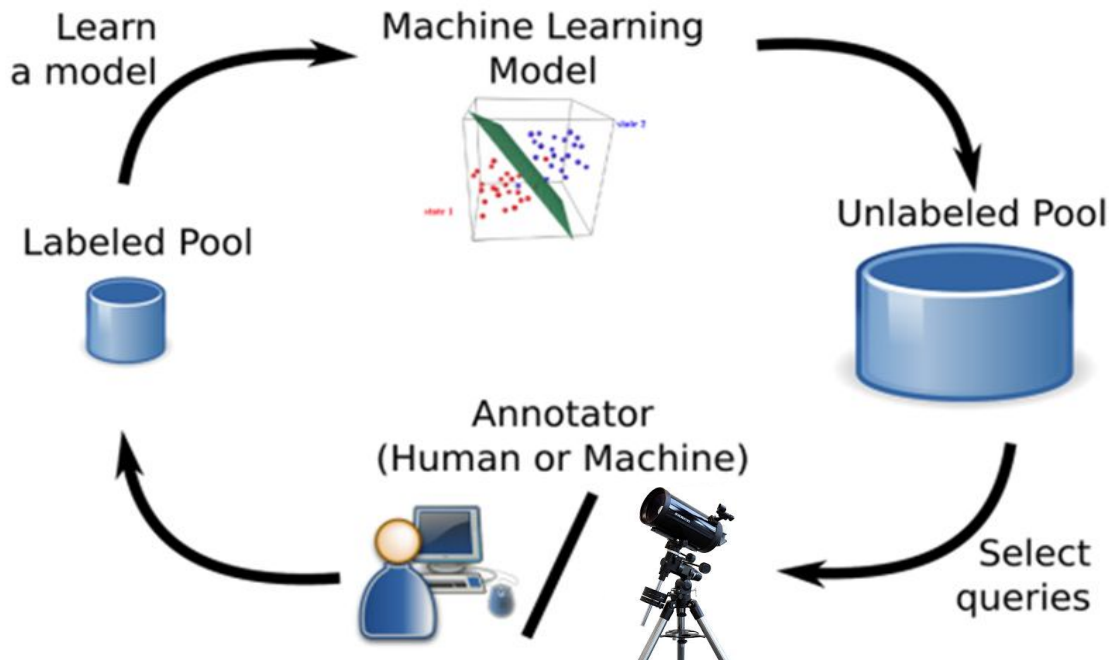
- 35 million candidate alerts
- 100 surviving selection cuts
- 6 followed-up by GRANDMA

*Extra imaging from
professional and
amateur
astronomers*



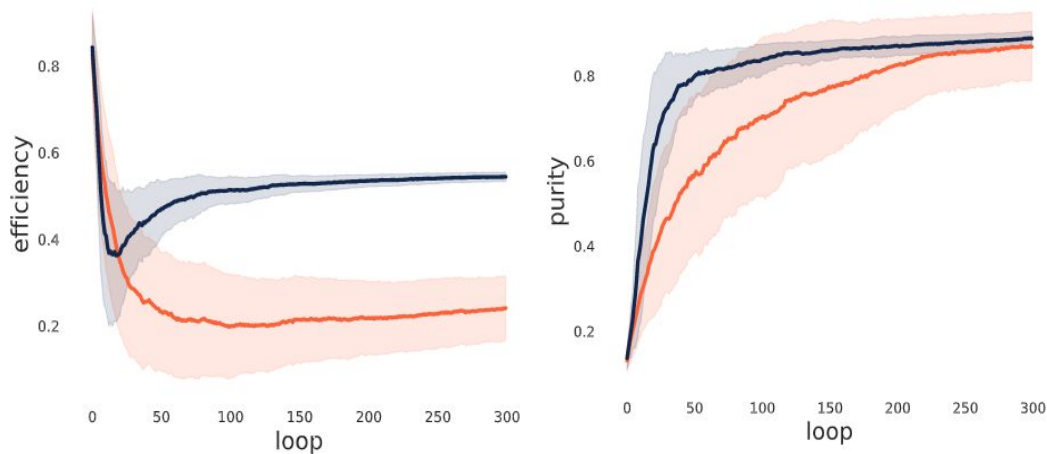
Case study: Early SN Ia classification

*Active Learning
Optimal experiment design*



Case study: Early SN Ia classification

— RandomSampling
— UncSampling



Results after 300 loops:

Training: 310 alerts
Testing: > 52 000 alerts

Choose training sample which lead to better results
and train a Random Forest classifier ...



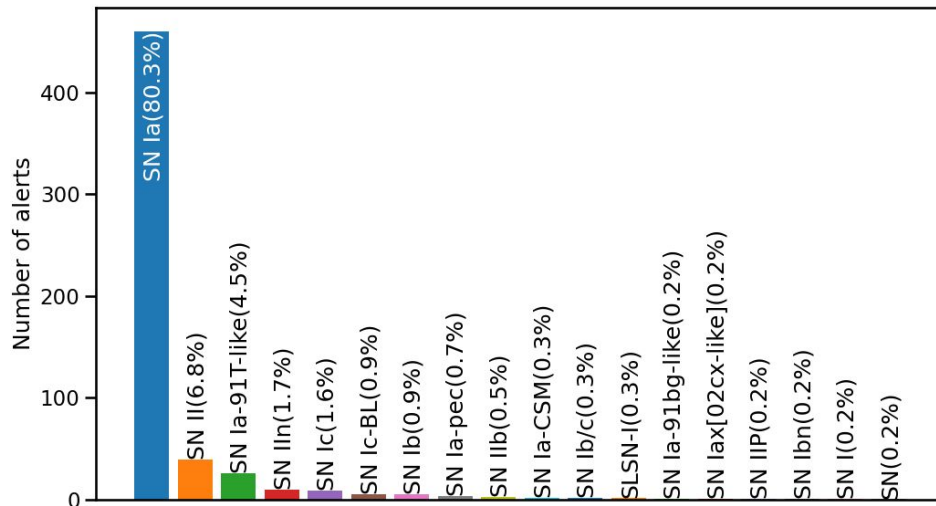
Trained ML
model



Case study: Early SN Ia classification

Fink Early SN Ia candidates reported to TNS from November/2020 - March/2022:

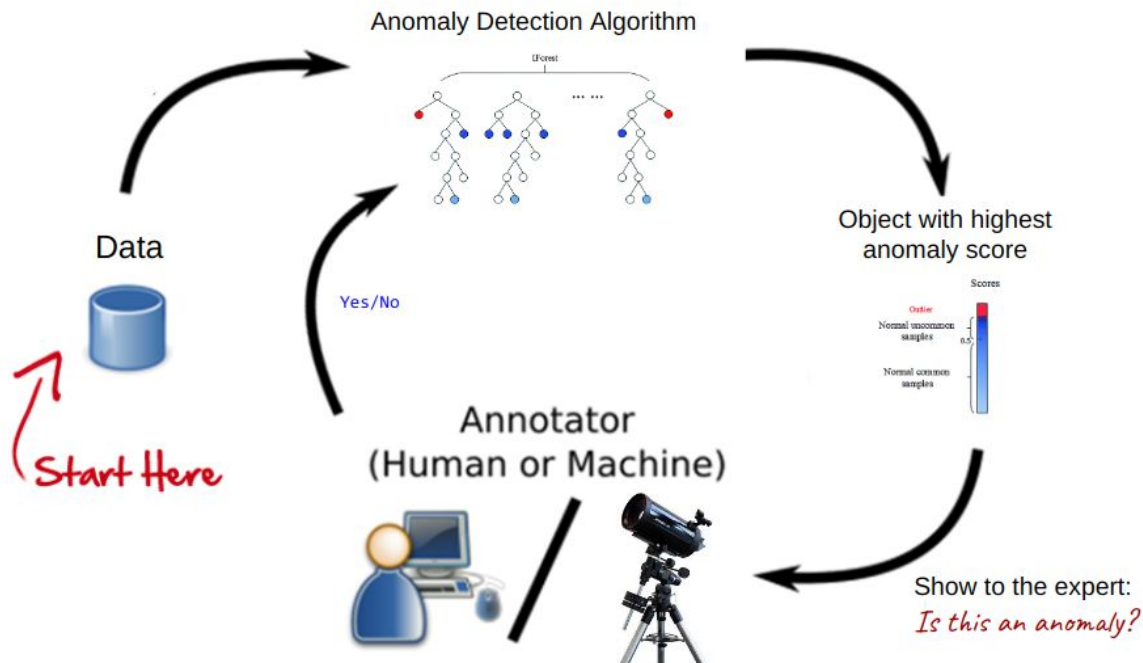
- 4 661 Early SN Ia candidates
- 573 spectroscopically classified
- Contaminants are mostly other SNe
 - 1 LBV



Plot by Julien Peloton (CNRS/IJCLab)



Active Anomaly Detection



Case study: Satellite tracks

Problem 1: they hide in plain sight. Labels must evolve

Not mega-constellations!

Problem 2: they move fast and may confuse difference image analysis

- Module to identify satellite glints
- 11.5 % all single-frame events
- 30% of those with real-bogus > 0.8
- 140 per night

