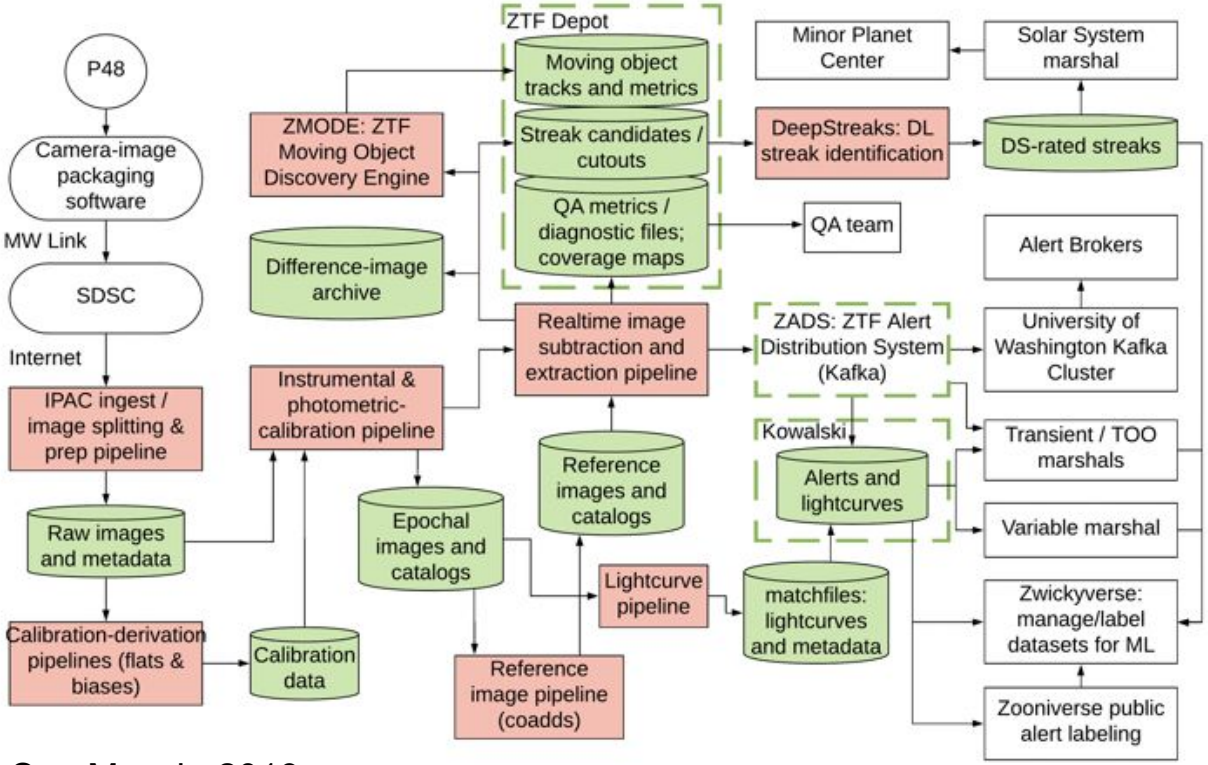
Fritz: science data platform for ZTF-II

Dmitry A. Duev

Research Scientist, Caltech

https://duev.space

ZTF-I: data/processing flow



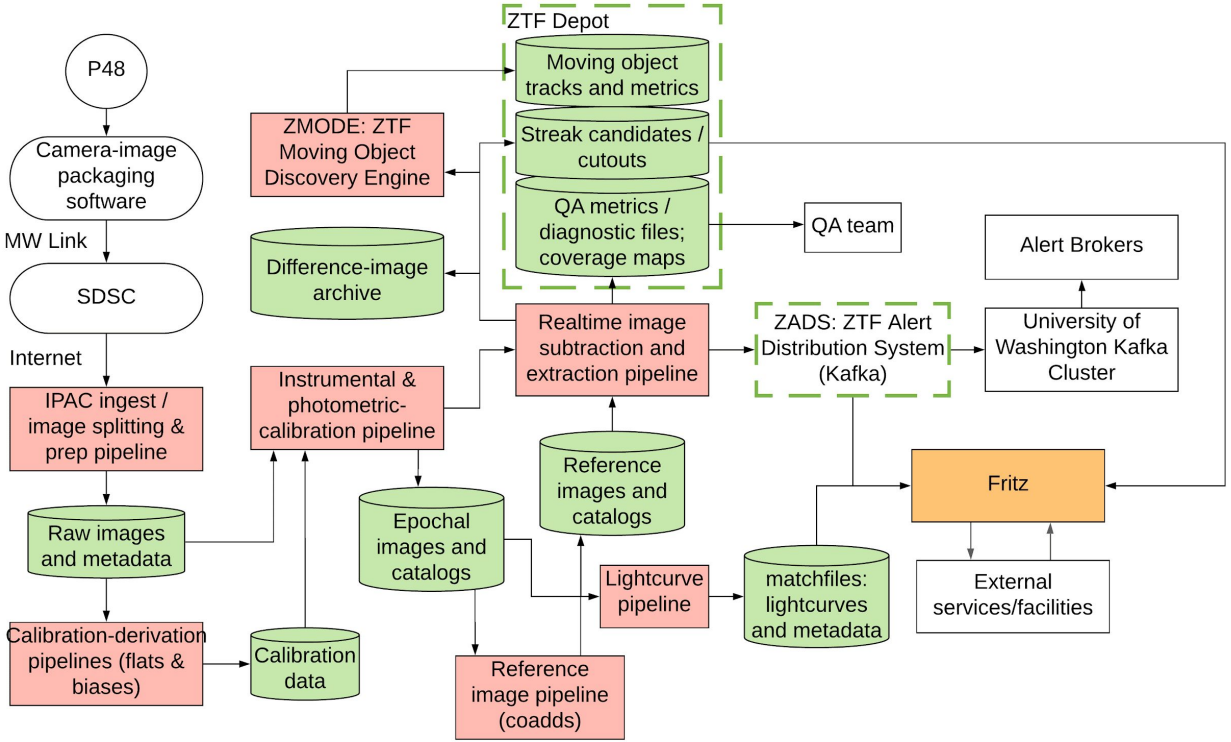
See Masci+ 2019

ZTF acts as a discovery engine. Discoveries are followed-up using a wide range of instruments (including ZTF itself)

Single night	8h40m
# on-sky exposures	~700
Raw image data	~1 TB
Real-time data products	~4 TB
# unvetted 5σ alerts	~10^5 - 10^6
# ML-vetted alerts	~10^3 - 10^5
# unvetted streaks	~10^4 - 10^6
# ML-vetted streaks	~10^2 - 10^3

Nominal survey	3 x 260 n
Volume of data products	~3.2 PB
Volume of ref images	~60 TB
# CCD quad ref images	~2.8×10^5
Volume of matchfiles	~50 TB
# matchfiles	~2.8×10^5
# single-epoch PSF-fit source measurements	~800 B
# single-epoch aperture source measurements	~230 B

ZTF-II: data/processing flow



ZTF acts as a discovery engine. Discoveries are followed-up using a wide range of instruments (including ZTF itself)

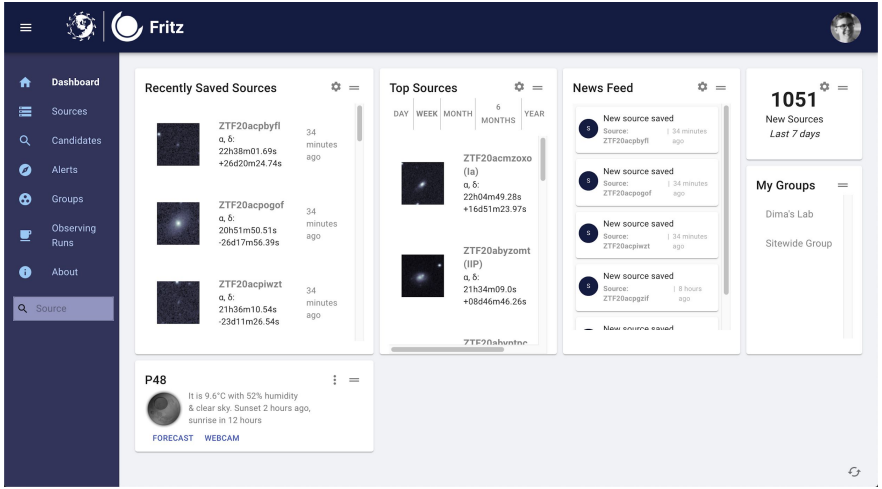
Single night	8h40m
# on-sky exposures	~700
Raw image data	~1 TB
Real-time data products	~4 TB
# unvetted 5σ alerts	~10^5 - 10^6
# ML-vetted alerts	~10^3 - 10^5
# unvetted streaks	~10^4 - 10^6
# ML-vetted streaks	~10^2 - 10^3

Nominal survey	3 x 260 n
Volume of data products	~3.2 PB
Volume of ref images	~60 TB
# CCD quad ref images	~2.8×10^5
Volume of matchfiles	~50 TB
# matchfiles	~2.8×10^5
# single-epoch PSF-fit source measurements	~800 B
# single-epoch aperture source measurements	~230 B

Fritz: science data platform for ZTF-II

- Multi-survey data archive and alert broker
- Marshal for transient, variable, and Solar system science cases
- Workhorse for ML applications/active learning: classification and labeling at scale
- Follow-up observation management: robotic and classical facilities

Initiated in Feb 2020
Beta up in Sep 2020
MVP live in Nov 2020



https://fritz-marshal.org/

Fritz: core dev team

Joshua Bloom

Michael Coughlin

Arien Crellin-Quick

Dmitry Duev

Daniel Goldstein

Matthew Graham

Mansi Kasliwal

Guy Nir

Kyung Min Shin

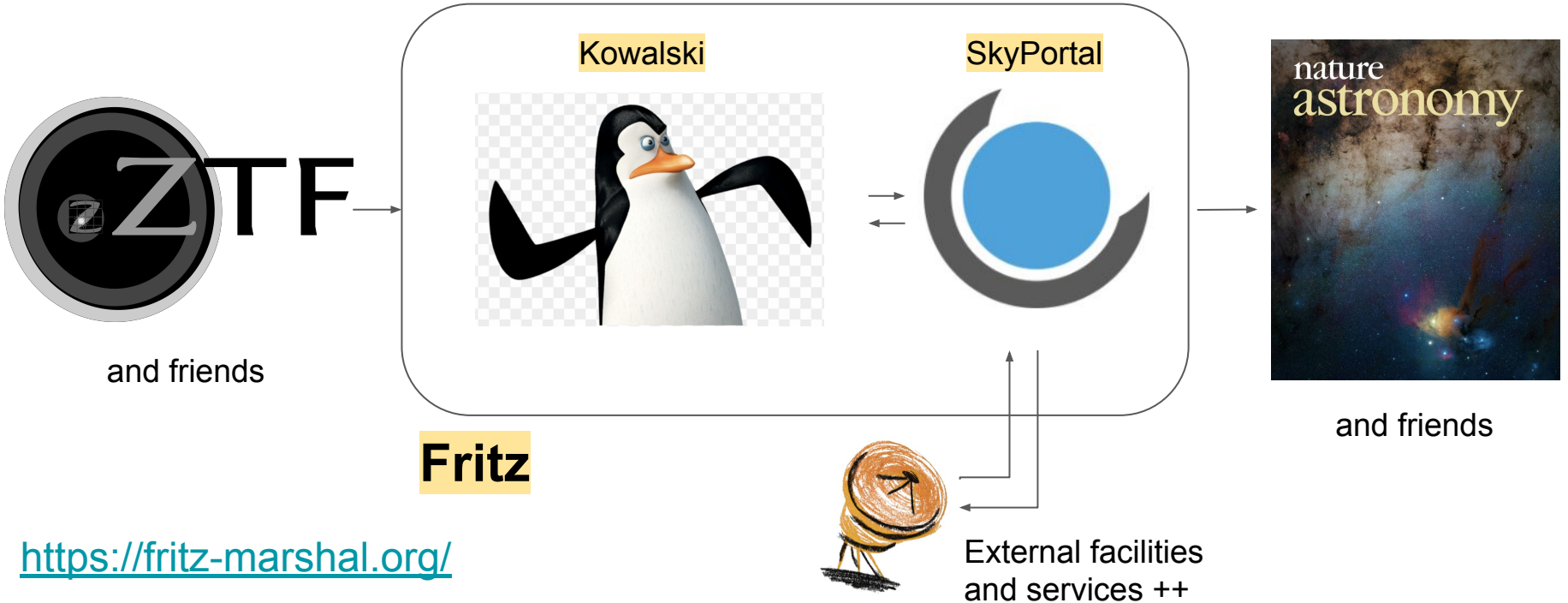
Leo Singer

Stéfan van der Walt

[Caltech + UC Berkeley ++]

Fritz: Schematic overview

Observe → Mine/Discover → Study and Characterize → Profit!



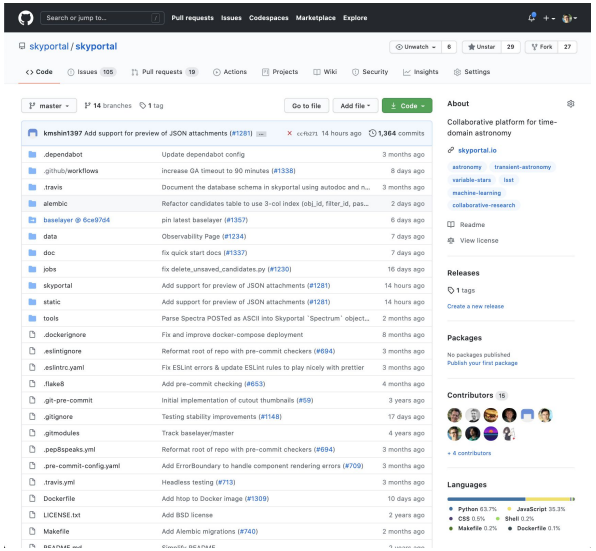
https://fritz-marshal.org/

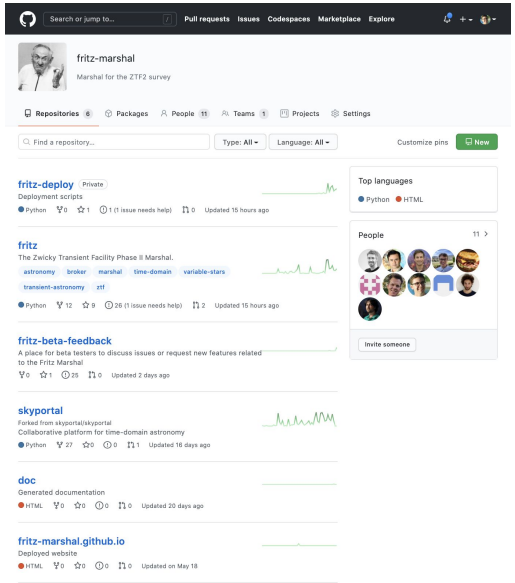
Fritz: features

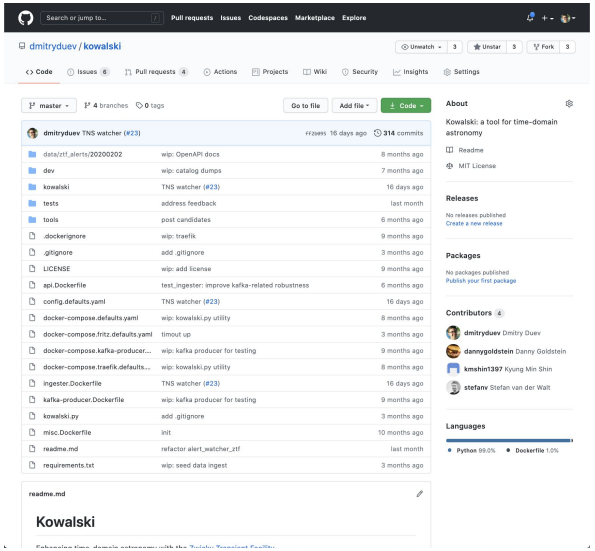
- Open source (free to use, modify, and distribute)
- API-first system: rich APIs for machine usage
- Powerful alert stream enhancement/filtering capabilities
- Extendible & scalable design: async Python backends, React/Redux frontend
- Fine-grained access control
- Authentication via OAuth
- Real-time Slack-like messaging, notifications
- Rich visualization capabilities
- Follow-up management
- Distributed computation via Dask
- Docker compose or Kubernetes deployment
- Well-tested, extensive docs, CI/CD

Fritz: features

- Open source (free to use, modify, and distribute)







Fritz: features

- API-first system: rich APIs for machine usage

Python

```
import requests

token = 'ea70a5f0-b321-43c6-96a1-b2de225e0339'

def api(method, endpoint, data=None):
    headers = {'Authorization': f'token {token}'}
    response = requests.request(method, endpoint, json=data, headers=headers)
    return response

response = api('GET', 'http://localhost:5000/api/sysinfo')

print(f'HTTP code: {response.status_code}, {response.reason}')
if response.status_code in (200, 400):
    print(f'JSON response: {response.json()}')
```

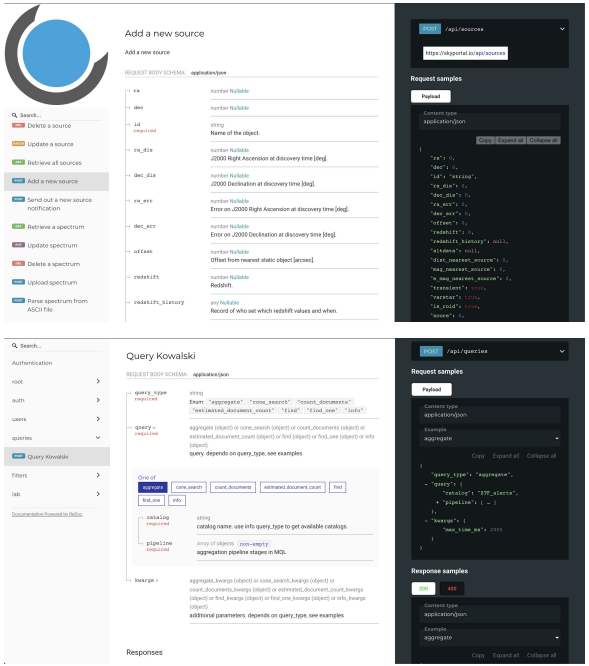
Command line (curl)

```
curl -s -H 'Authorization: token ea70a5f0-b321-43c6-96a1-b2de225e0339' http://
```

Response

In the above examples, the SkyPortal server is located at `http://localhost:5000`. In case of success, the HTTP response is 200:

```
HTTP code: 200, OK
JSON response: {'status': 'success', 'data': {}, 'version': '0.9.dev0+git2020
```



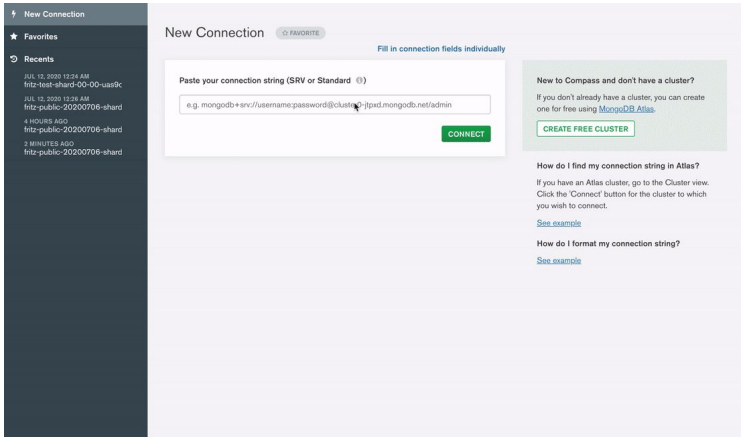
OpenAPI specs:

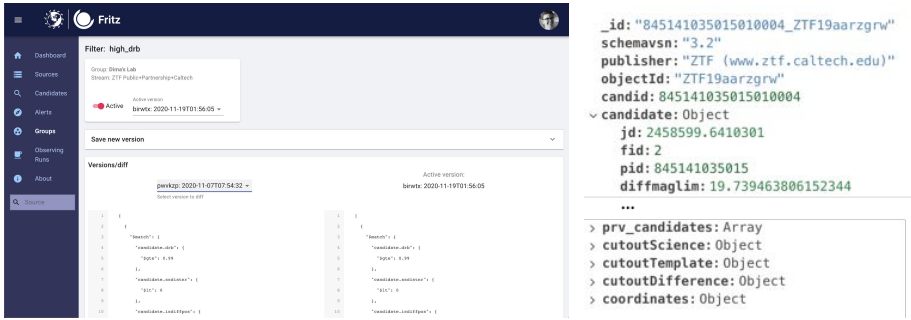
https://skyportal.io/docs/api.html

https://kowalski.caltech.edu/docs/api

Fritz: features

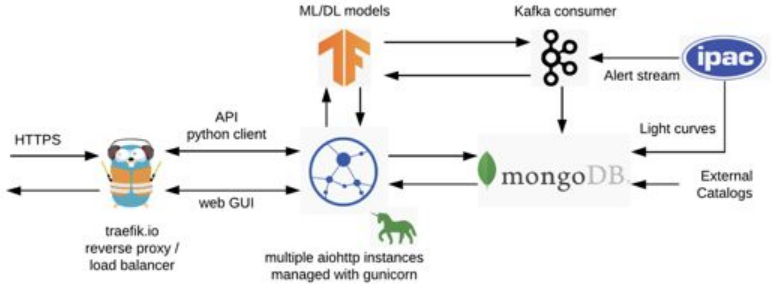
- Fast, robust access multi-survey data archive
  - ZTF alert stream
  - ZTF light curves + features + SCoPe classifications
  - >30 external catalogs
  - Multiple ML models
  - Cross-matches
  - ~100M queries/day typical load
- Powerful alert stream enhancement & filtering capabilities
  - MongoDB aggregation pipelines
  - Public alert DBs for filter design/debugging
  - Filtering enhanced data
  - Automated checks, no filter code audit
  - [Can post results from external brokers]

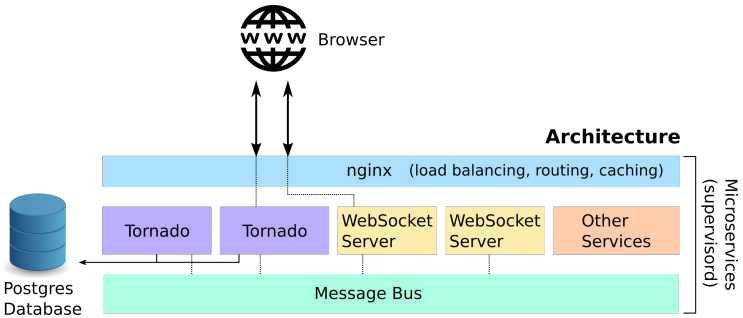


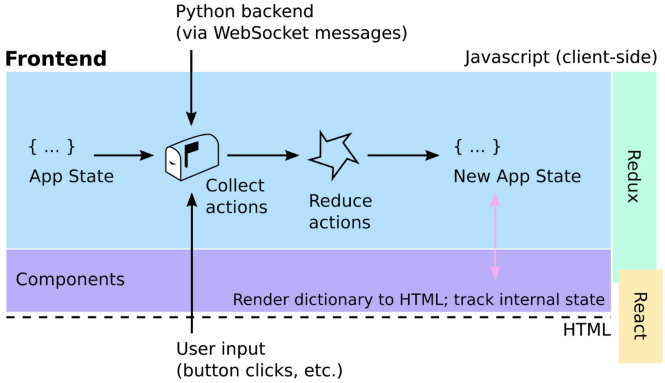


Fritz: features

- Extendible & scalable design: async Python backends, React/Redux frontend

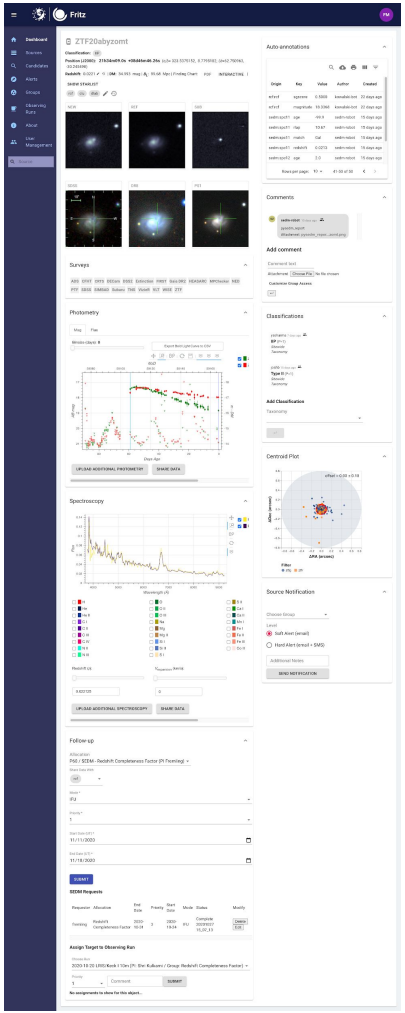


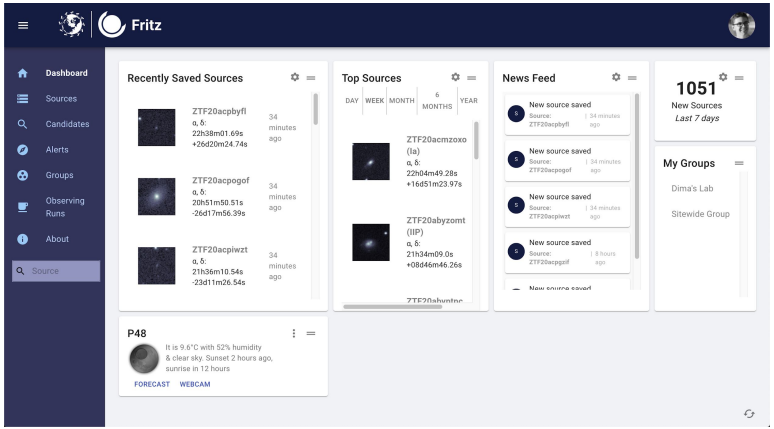


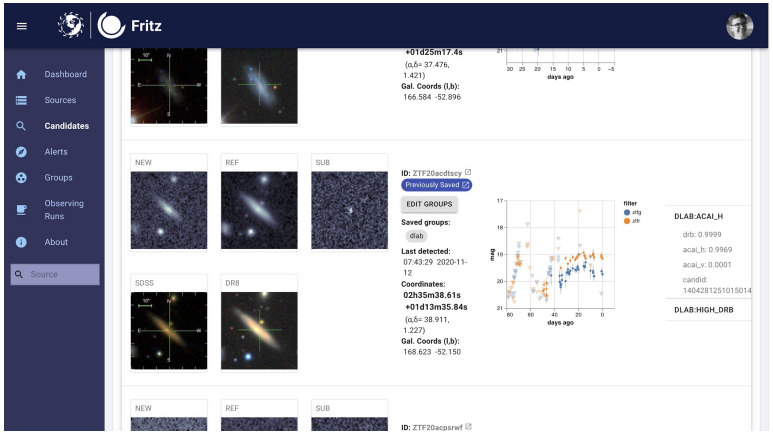


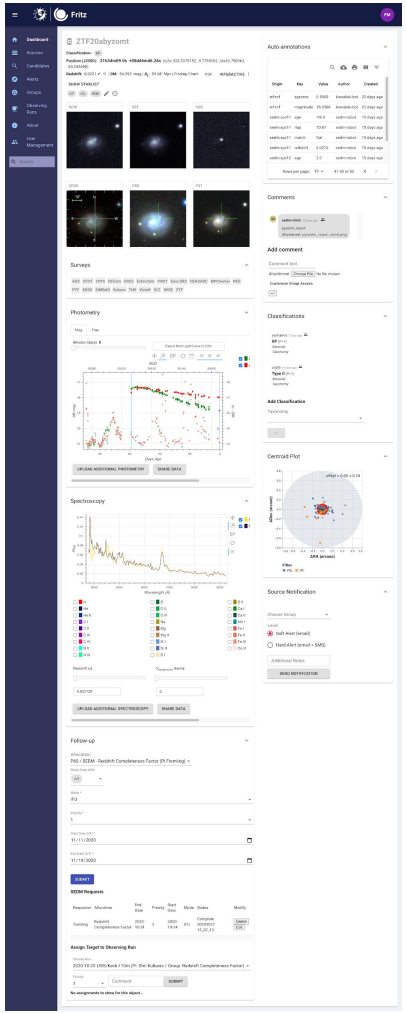
Fritz: features

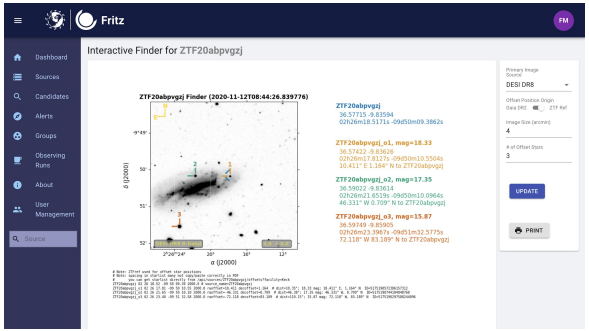
- Fine-grained access control
  - ACLs, Roles
  - Groups (>30 as of this morning; ~200 users)
- Authentication via OAuth
- Real-time Slack-like messaging, notifications (email/text)
- Rich visualization capabilities
- Follow-up observations management (TOM)
  - Both robotic and classical facilities (SEDM end-to-end)
- Distributed computation via Dask
- Docker compose or Kubernetes deployment
  - Everything containerized / versioned
  - Production deployment: Local + GCP
- Well-tested, extensive docs, CI/CD, staging

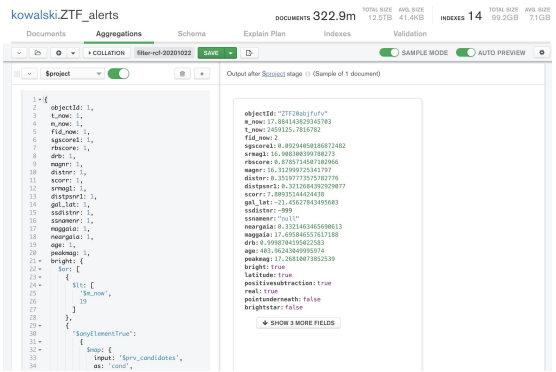



Fritz
Dashboard
Sources
Candidates
Alerts
Groups
Observing Runs
About
Source
Recently Saved Sources
ZTF20acpbyfl
α, δ: 22h38m01.69s +26d20m24.74s
34 minutes ago
ZTF20acpogof
α, δ: 20h51m50.51s -26d17m56.39s
34 minutes ago
ZTF20acpiwzt
α, δ: 21h36m10.54s -23d11m26.54s
34 minutes ago
Top Sources
DAY WEEK MONTH 6 MONTHS YEAR
ZTF20acmzoxo (Ia)
α, δ: 22h04m49.28s +16d51m23.97s
ZTF20abyzomt (IIP)
α, δ: 21h34m09.0s +08d46m46.26s
News Feed
New source saved
Source: ZTF20acpbyfl
34 minutes ago
New source saved
Source: ZTF20acpogof
34 minutes ago
New source saved
Source: ZTF20acpiwzt
34 minutes ago
New source saved
Source: ZTF20acpgzil
8 hours ago
1051
New Sources
Last 7 days
My Groups
Dima's Lab
Sitewide Group
P48
It is 9.6°C with 52% humidity & clear sky. Sunset 2 hours ago, sunrise in 12 hours
FORECAST WEBCAM


Fritz
Dashboard
Sources
Candidates
Alerts
Groups
Observing Runs
About
Source
NEW
REF
SUB
SDSS
DR8
ID: ZTF20acdtscy
Previously Saved
EDIT GROUPS
Saved groups:
dlab
Last detected:
07:43:29 2020-11-12
Coordinates:
02h35m38.61s +01d13m35.84s
(α,δ= 38.911, 1.227)
Gal. Coords (l,b):
168.623 -52.150
+01d25m17.4s
(α,δ= 37.476, 1.421)
Gal. Coords (l,b):
166.584 -52.896
days ago
filter
DLAB:ACAI_H
drb: 0.9999
acai_h: 0.9969
acai_v: 0.0001
candid: 1404281251015014
DLAB:HIGH_DRB
ID: ZTF20acpsrwf


Fritz
ZTF20abyzomt
Auto annotations
Comments
Add comment
Surveys
Photometry
Classifications
Add Classification
Centroid Plot
Spectroscopy
Source Notification
Follow-up
UPLOAD ADDITIONAL PHOTOMETRY
SHARE DATA
UPLOAD ADDITIONAL SPECTROSCOPY
SUBMIT


Fritz
Interactive Finder for ZTF20abpvgzj
ZTF20abpvgzj Finder (2020-11-12T08:44:26.839776)
ZTF20abpvgzj
36.57715 -9.83594
02h26m18.5171s -09d50m09.3862s
ZTF20abpvgzj_o1, mag=18.33
36.57422 -9.83626
02h26m17.8127s -09d50m10.5504s
10.411" E 1.164" N to ZTF20abpvgzj
ZTF20abpvgzj_o2, mag=17.35
36.59022 -9.83614
02h26m21.6519s -09d50m10.0964s
46.331" W 0.709" N to ZTF20abpvgzj
ZTF20abpvgzj_o3, mag=15.87
36.59749 -9.85905
02h26m23.3967s -09d51m32.5775s
72.118" W 83.189" N to ZTF20abpvgzj
Primary Image Source
DESI DR8
Offset Position Origin
Gaia DR2
ZTF Ref
Image Size (arcmin)
4
# of Offset Stars
3
UPDATE
PRINT


kowalski.ZTF_alerts
DOCUMENTS 322.9m
TOTAL SIZE 12.5TB
AVG. SIZE 41.4KB
INDEXES 14
TOTAL SIZE 99.2GB
AVG. SIZE 7.1GB
Documents
Aggregations
Schema
Explain Plan
Indexes
Validation
+ COLLATION
filter-rcf-20201022
SAVE
SAMPLE MODE
AUTO PREVIEW
$project
Output after $project stage (Sample of 1 document)
{
objectId: 1,
t_now: 1,
m_now: 1,
fid_now: 1,
fid_now: 2
sgscore1: 1,
rbscore: 1,
drb: 1,
magnr: 1,
magnr: 16.312999725341797
distnr: 0.351977735757827776
distpsnr1: 0.321684992929077
scorr: 7.789351344224438
gal_lat: -21.456278434995603
ssdistnr: -999
ssnamenr: "null"
neargaia: 0.3321463465698613
maggaia: 17.695846557617188
drb: 0.9998704195022583
age: 403.96243049995974
peakmag: 17.268100738525339
bright: true
latitude: true
positivesubtraction: true
real: true
pointunderneath: false
brightstar: false
SHOW 3 MORE FIELDS
objectId: "ZTF20abjfufv"
$lt: [
"$m_now",
19
]
},
{
"$anyElementTrue":
{
$map: {
input: '$prv_candidates',
as: 'cand',

DEMO