



Sherlock

A contextual classifier for transient surveys

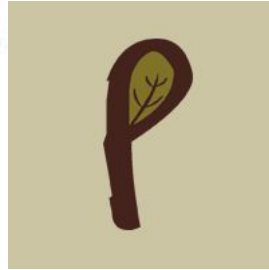
Dave Young, Stephen Smartt, Ken Smith (QUB)



Sherlock

Executive Summary

Sherlock is a boosted decision tree algorithm that mines a library of historical and on-going astronomical survey data in an attempt to predict the nature of the object based on the resulting crossmatched associations found.



Overview

- The Purpose of Lasair Annotations (e.g. Sherlock)
- Mechanics of Sherlock
- Looking towards Rubin's LSST & the future of Sherlock.

The Purpose of Lasair ‘Annotations’



1. What is your favourite flavour of transient?
2. Define a checklist of attribute values that would allow you to uniquely select out your selected flavour
3. Annotate all transients with any missing attribute
4. Filter/search/sort you favourite transient flavours from streams, API, webpage-search

Lasair API 

} **temporal**: lightcurve metrics
spatial: environment/context
multi-messenger: GW, neutrinos,





ZTF23abbabr
124.201240, 48.816158



Discovery Date: 2023-09-05 11:55:25 UTC
Discovery MJD: 60192.50
Disc r-Mag: 20.23±0.24

Latest Date: 2023-09-07 12:04:16 UTC
Latest MJD: 60194.50
Latest r-Mag: 19.75±0.17



Transient Name Server
AT2023rpf

The transient was discovered on 5th September 2023 at 11:55:40 (MJD 60192.50) by ZTF as ZTF23abbabr with a discovery magnitude of $r = 20.23$.



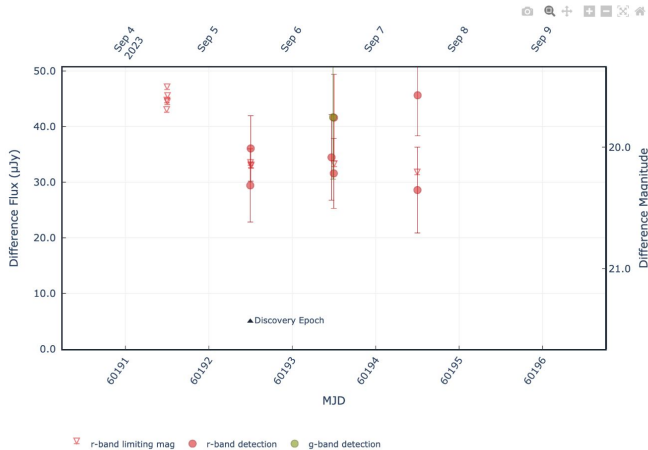
Sherlock Contextual Classification
Prediction: **Supernova**

The transient is possibly associated with [SDSSJ081648.49%2B484857.3](#); an $r=20.07$ mag galaxy found in the NED/SDSS/PS1 catalogues. Its located 1.05" N, 2.27" W from the galaxy centre.



Difference Image Lightcurve

Access the data from this plot in the ZTF Alert Packet Data table below.



Learn how to [interact with this plot](#).

Every transient in Lasair now has a Sherlock annotation giving a top level predicted classification alongside some extra metadata.

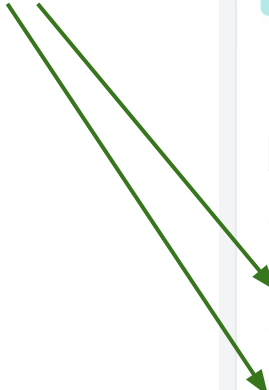




Annotators

Annotators are codes that run on the Lasair transient stream, compute some value added feature(s) for the transients and inject these values back into the stream for other users to use in their filters. Some annotators are written by the Lasair Team but others are submitted by a Lasair users. Please get in contact via [Community](#) if you have an idea for an annotator.

2nd Generation Annotators



Public Gallery

Annotators submitted to the public gallery by the Lasair Team or other Lasair users.

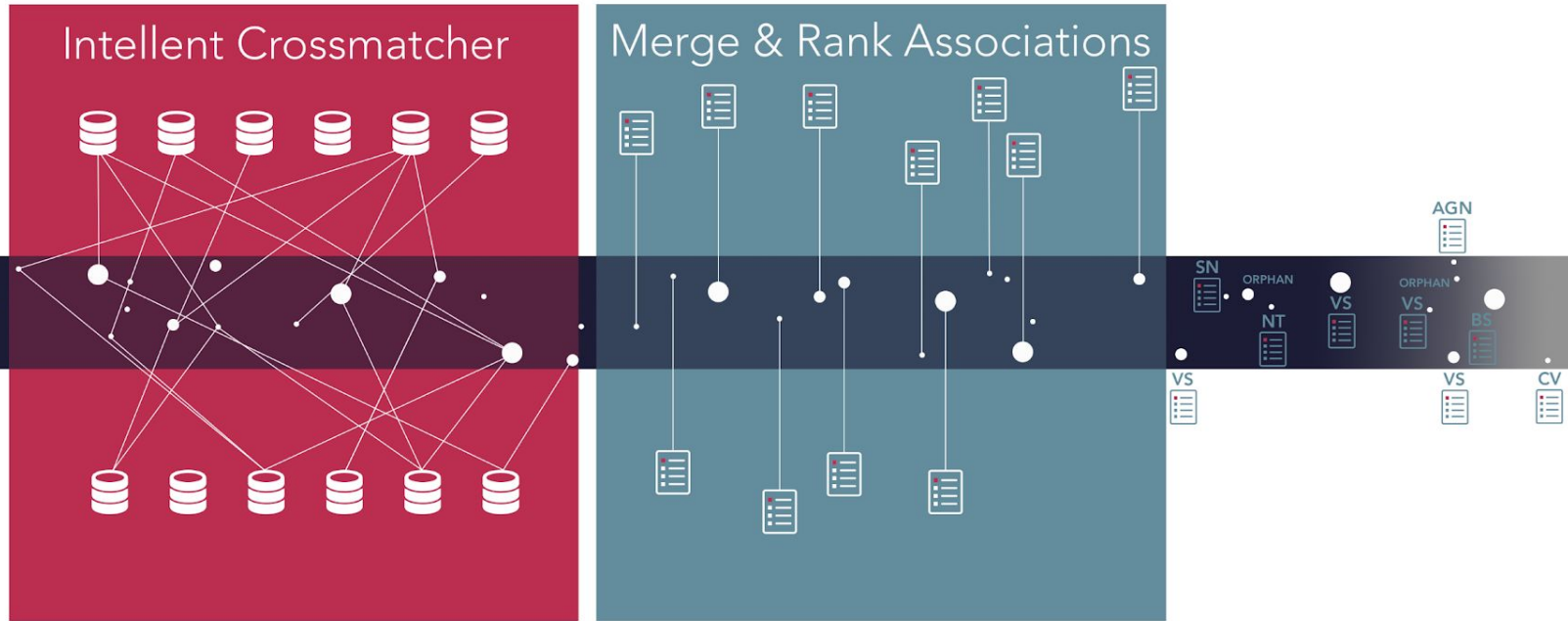
Topic	Owner	Description	Count
alerce_ic	Admin Admin	The ALeRCE light cu ... more	199,674
alerce_stamp	Admin Admin	The Alerce Stamp CI ... more	3,479
fastfinder	Michael Fulton	Fastfinder is an early-time, fast transient alerting system ... more	12,118
Fink	Admin Admin	Fink is a LSST Community ... more	17,180
NEEDLE	Xinyue Sheng	Finding TDEs and SLSN from image stamps	177
slowSN	Philip Wiseman	Mining Lasair for other examples of AT2021lwx also known as ... more	7
test	Roy Williams	Testing	11

100 rows per page

Showing 1 to 7 of 7 rows

<https://lasair-ztf.lsst.ac.uk/annotators/>

The Mechanics of Sherlock



Catalogue Library



Source-agnostic(ish) all-sky surveys:

- Gaia DR2,
- PanSTARRS DR1
- GSC v2.3,
- SDSS DR12 PhotoObjAll Table,
- SDSS DR12 SpecObjAll Table,
- 2MASS catalogues
- ...

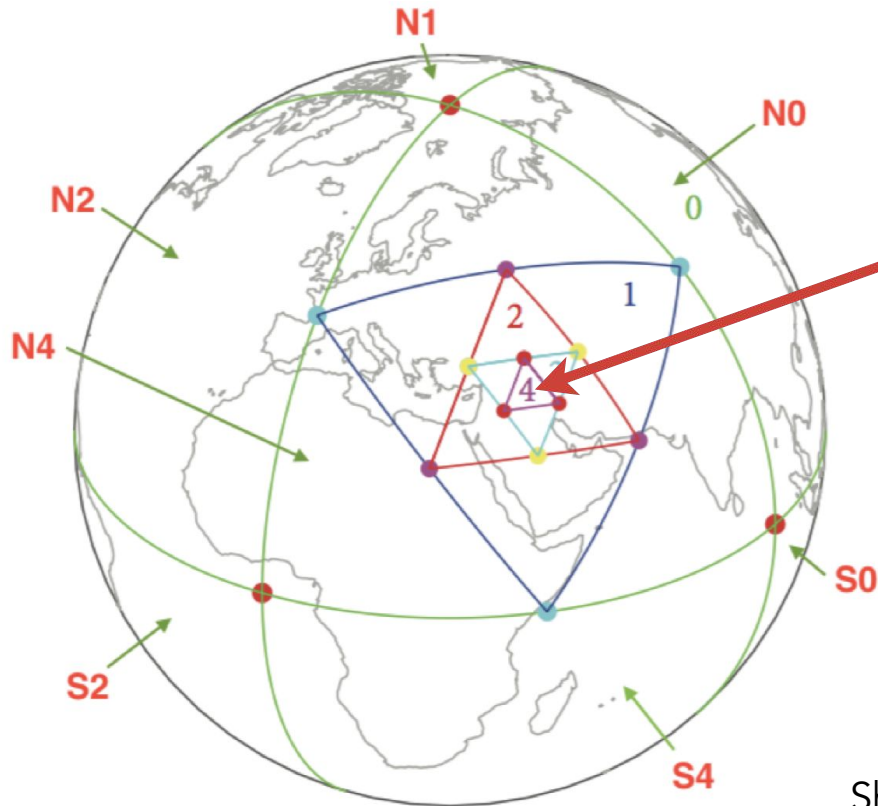
Smaller source-specific catalogues:

- Million Quasars Catalog v5.2,
- Veron-Cett AGN Catalogue v13,
- Downes Catalog of CVs,
- Ritter Cataclysmic Binaries Catalog v7.21,
- LASr galaxy catalogue
- NED-D Galaxy Catalogue v13.1
- ...

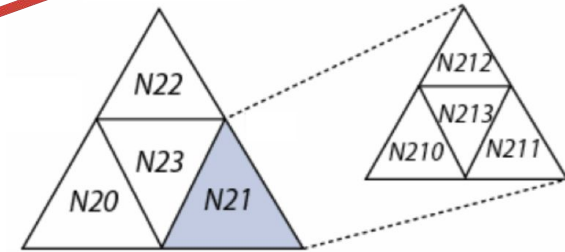
Remote query and caching of NASA/IPAC Extragalactic Database (NED)

4.1 TB MySQL Database. Multiple copies in QUB and ROE

Hierarchical Triangular Mesh (HTM)



N43333



HTM was first proposed in Kunszt 2000 as a solution to:

1. *indexing* the surface of a sphere.
2. provide the ability to query *positions* and *areas* on the surface of the sphere.

Sherlock tables indexed with HTM-levels 10, 13 and 16

Plain Text Algorithms

```
search algorithm:
GSC star 1:
  database table: tcs_view_star_guide_star_catalogue_v2_3
  mag column: V
  bright:
    mag limit: 16.
    angular radius arcsec: 100.0
    synonym: VS
    association: BS
  faint:
    mag limit: 19.5
    angular radius arcsec: 2.0
    annotation: SN
  general:
    angular radius arcsec: 0.5
    synonym: VS
```

```
GSC star 2:
  database table: tcs_view_star_guide_star_catalogue_v2_3
  mag column: B
  bright:
    mag limit: 16.
    angular radius arcsec: 100.0
    synonym: VS
    association: BS
  faint:
    mag limit: 19.5
    angular radius arcsec: 2.0
    annotation: SN
  general:
    angular radius arcsec: 0.5
    synonym: VS
```

Search Module Parameters

- *angular separation crossmatch radius*
- *physical separation crossmatch radius*
- *source magnitude filtering*
- *magnitude dependent search radii for bright stars and galaxies*

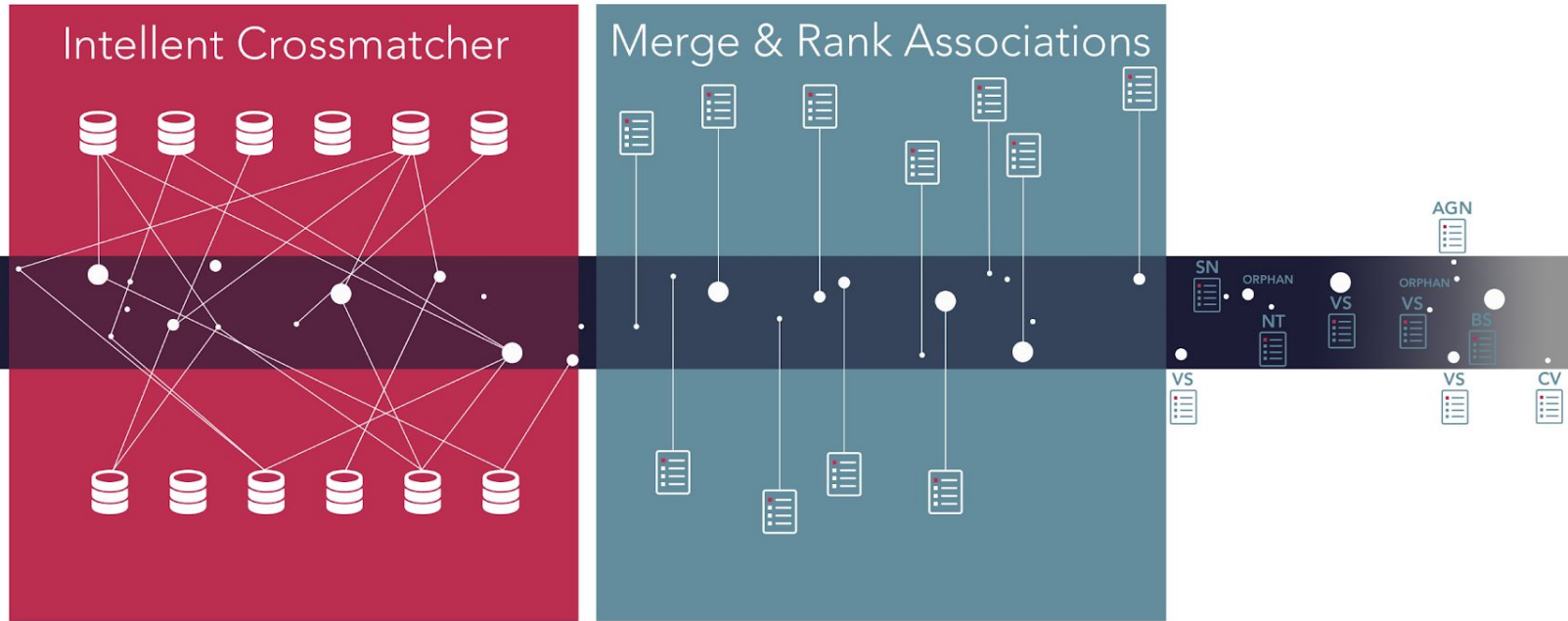
Classifications & Reliabilities

- *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*

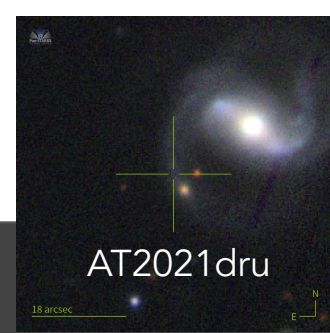
Annotations: Contextual Classifications

1. **Variable Star (VS)** if the transient lies within the synonym radius of a catalogued point-source,
2. **Cataclysmic Variable (CV)** if the transient lies within the synonym radius of a catalogued CV,
3. **Bright Star (BS)** if the transient is not matched against the synonym radius of a star but is associated within the magnitude-dependent association radius,
4. **Nuclear Transients (NT)** if the transient falls within the synonym radius of the core of a resolved galaxy,
5. **Supernova (SN)** if the transient is not classified as an NT but is found within the magnitude-, morphology- or distance-dependant association radius of a galaxy, or
6. **Active Galactic Nucleus (AGN)** if the transient is matched against the synonym radius of a known AGN.
7. **Orphan** if the transient fails to be matched against any catalogued source.

The Mechanics of Sherlock



Command-Line Usage



```
> sherlock -N match -- 279.85096 -23.60341
```

The transient is possibly associated with [2MASXJ12244364-0445584](#); a J=12.53 mag galaxy found in the NED/2MASS/PS1 catalogues. It's located 8.62" S, 15.21" E (9.0 Kpc) from the galaxy centre. A host $z=0.026$ implies a $m - M = 35.27$.

Transient's Predicted Classification: SN

Suggested Associations:

association type	rank	rankScore	catalogue table name	catalogue object id	catalogue name
SN	1	1952.45	NED/2MASS/PS1	2MASXJ12244364-0445584	galaxy
SN		1952.45	2MASS XSC	12244364-0445584	galaxy
SN		1952.45	NED	2MASX J12244364-0445584	galaxy
BS		1952.45	PanSTARRS DR1	102281861819500963	galaxy
SN	2	2005.00	PS1	102271861849167994	galaxy
SN		2005.00	PanSTARRS DR1	102271861849167994	galaxy
SN	3	2005.00	PS1	102271861855876936	galaxy
SN		2005.00	PanSTARRS DR1	102271861855876936	galaxy

Annotations: Bonus Extras

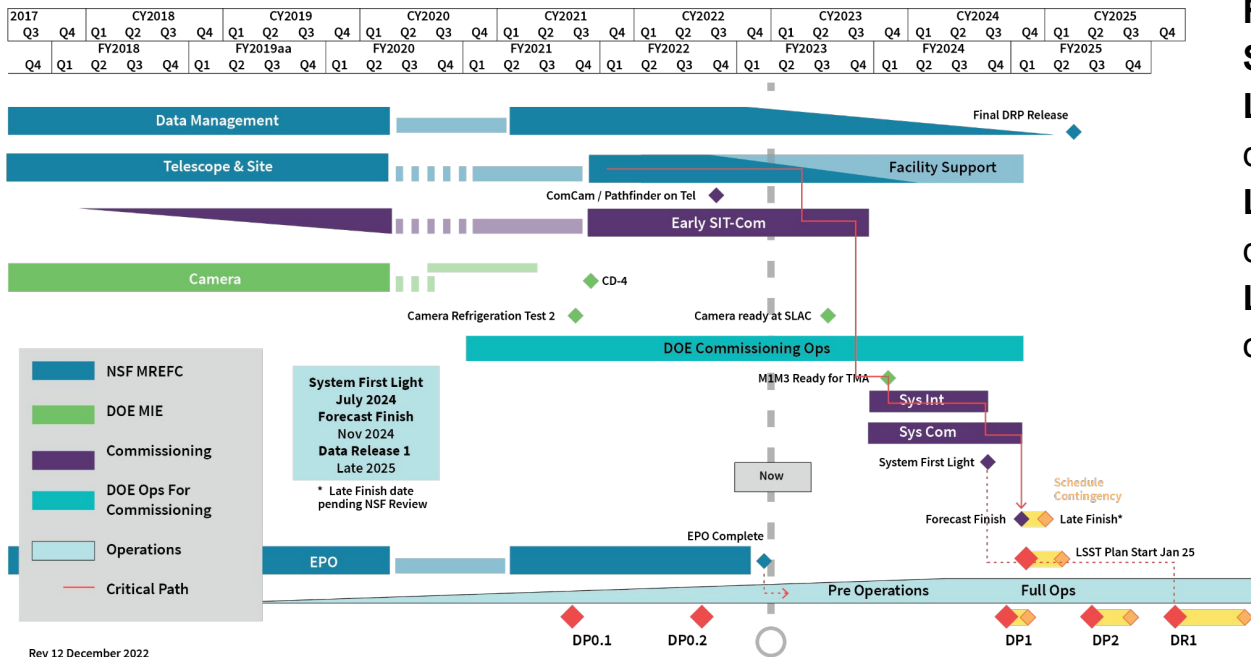
- Synonymous/Associated source IDs
- Multi-wavelength magnitudes
- Host redshifts and direct distances (if available)
- Host morphologies
- Host colours
- ...

The greater the number, quality and diversity of annotations hosted by Lasair, the more powerful its search and filtering engines become and the more creative our end users can be.

Big Fat Caveat

The predicted nature of any transient source Sherlock provides is only ever as good as the underlying data it has used to provide that prediction.

Sherlock and LSST



First Light: Apr 2025

Survey Operations Start: ~Nov 2025

LSST DR1: Nov 2026 (first 6 months of data)

LSST DR2: Feb-Nov 2027 (first year of data)

LSST DR3: Feb-Nov 2028 (first 2 years of data)

- <https://www.lsst.org/about/timeline>
- <https://www.lsst.org/scientists/early-science>
- <https://rtn-011.lsst.io/>

The Future of Sherlock

- Need to include methods to identify high-proper motion stars in Sherlock (would need a detection epoch and position)
- Many other catalogue can and will be add to the Sherlock database.
- Improve the Merge and Ranking algorithm (with ML)
- Continually feedback (from users) and improvement of the algorithm