Enabling Citizen Science with Rubin Observatory Alert Stream Brokers

Meg Schwamb, Chris Lintott, Ken Smith, Jamie Robinson, Roy Williams, Lucy Fortson, Michael Fulton, Grant Miller, Rafia Omer, Stephen Smartt, and Darryl Wright

Corresponding Authors: Meg Schwamb (m.schwamb@qub.ac.uk), Chris Lintott (chris.lintott@physics.ox.ac.uk**)**

Introduction

The Vera C. Rubin Observatory's Legacy Survey of Space and Time (LSST; Ivezíc et al. 2019) will radically transform time-domain astronomy. Within approximately 60s of closing its camera shutter, the Rubin Observatory's automated data reduction routines will process and search the resulting image, covering 9.62 square-degrees, for astronomical transients. During the ten years of LSST, these pipelines will send 10,000 transient alerts per observation, with ten million LSST alerts expected to be generated by the end of each night. Rapid multi-wavelength follow-up observations would allow for probing the drivers/progenitors of the rare and interesting transients observed by LSST.

Sifting through this firehose of data to find the interesting astrophysical phenomena and Solar System transients is one of the key challenges of the LSST erra (see the LSST Science Book, LSST Science Collaboration 2019, for further details). Part of the solution will be to automatically sort and filter the LSST alert stream by a handful of software 'brokers' in order to identify interesting astronomical phenomena for astronomers to rapidly follow-up and study. Another answer to the solution will involve partnering the automated routines with human pattern recognition with online citizen science. In this report, we highlight a potential avenue for enabling a citizen science approach with the LSST alert stream through alert brokers.

The Zooniverse and LSST Citizen Science

The Zooniverse[1] is the largest platform for online citizen science projects,with over 200 projects designed and built with its Project Builder Platform[2], which provides web tools to enable researchers to quickly develop online citizen science projects for a suite of different types of tasks. The Zooniverse has partnered with the Rubin Observatory as their official citizen science partner. LSST Zooniverse projects can be built with proprietary data via the Rubin Observatory US data access center and a dedicated module with the Project Builder Platform. Zooniverse projects can also be built using alert stream data provided via the alert brokers, but it is up to

[1] https://www.zooniverse.org
[2] https://www.zooniverse.org/lab

the researcher to currently handle how to get the data sent by the different brokers into a format suitable for the Zooniverse platform.

Supernovae Hunters[3] (Wright et al 2017) is a present example of the kinds of potential future LSST citizen science projects that may develop. Supernovae Hunters, as shown in Figure 1, is enlisting the public to review LSST precursor data, Pan-STARRS (Panoramic Survey Telescope and Rapid Response System; Kaiser et al. 2010) images, to help search for supernovae by filtering the real supernovae detections from the large number of false positives identified by the automatic routines.

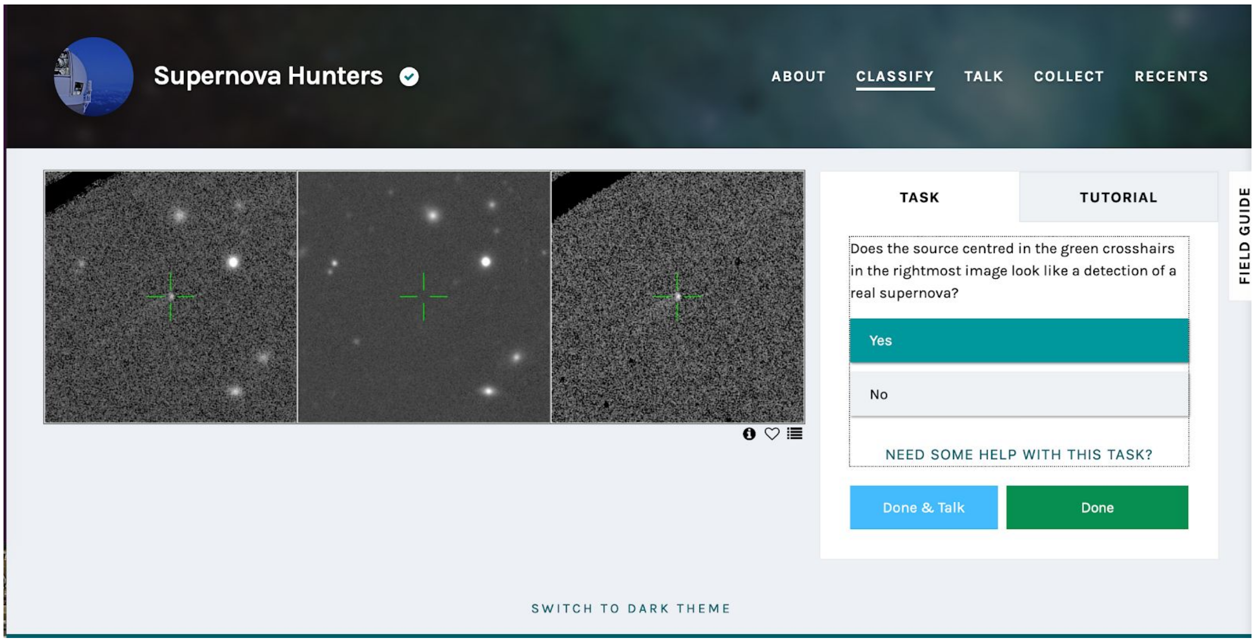

Figure 1: Classification interface from the Supernovae Hunters online citizen science project which is searching for supernovae in potential transients identified in the Pan-STARRS data. The left image is a recent image from the telescope with the candidate supernovae location highlighted with the green crosshairs, The center frame is a reference/template image of the same region taken previously. On the right is the difference image made by subtracting the reference image from the left frame.

Working with a Broker

It is anticipated that the world public alert stream from LSST (Bellm et al. 2020) will be parsed by brokers and made available to the research community worldwide. Current LSST alert broker concepts include ALeRCE[4] (Automatic Learning for the Rapid Classification of Events), AMPEL[5] (Alert Management, Photometry and Evaluation of Lightcurves; Norden et al. 2019), ANTARES[6]

[3] https://www.zooniverse.org/projects/dwright04/supernova-hunters
[4] http://alerce.science/
[5] https://github.com/AmpelProject/Ampel-contrib-sample
[6] https://antares.noao.edu/

(Arizona-NOAO Temporal Analysis and Response to Events System; Saha et al. 2016, Narayan et al. 2018), Fink[7], and Lasair[8] (Smith et al. 2019).

We use the example of Lasair, which currently serves Zwicky Transient Facility (ZTF) alerts (Patterson et al. 2019), to illustrate how this will work (see Figure 2). Lasair annotates the alert stream (using a module known as Sherlock[9]) with contextual classifications, which can then, along with metadata provided in the alert stream itself be used to filter the candidate alerts. The main Lasair site provides astronomers with the ability to search for alerts, but also to produce an 'active search' which will update as new alerts matching the search criteria arrive. The results of such an active search can be broadcast as a kafka stream[10] of object IDs which will be the primary way we expect astronomers using Lasair to identify alerts which can be sent to a Zooniverse citizen science project.

In this scenario the majority of the connections between the Zooniverse and Lasair are setup outside of both systems with the researcher running their own computer or server to receive the desired set of ZTF/LSST alerts from Lasair, parse the information, and create subjects (images, videos, plots, audio files, etc reviewed by volunteers on citizen science projects), and upload the subjects to the Zooniverse platform. This arrangement ensures that the system is agnostic to the choice of technology, for example for machine learning, used by the researcher and allows for flexibility in approach that would not otherwise be possible.

[7] https://fink-broker.org/
[8] https://lasair.roe.ac.uk/
[9] https://qub-sherlock.readthedocs.io/en/stable/
[10] https://kafka.apache.org/

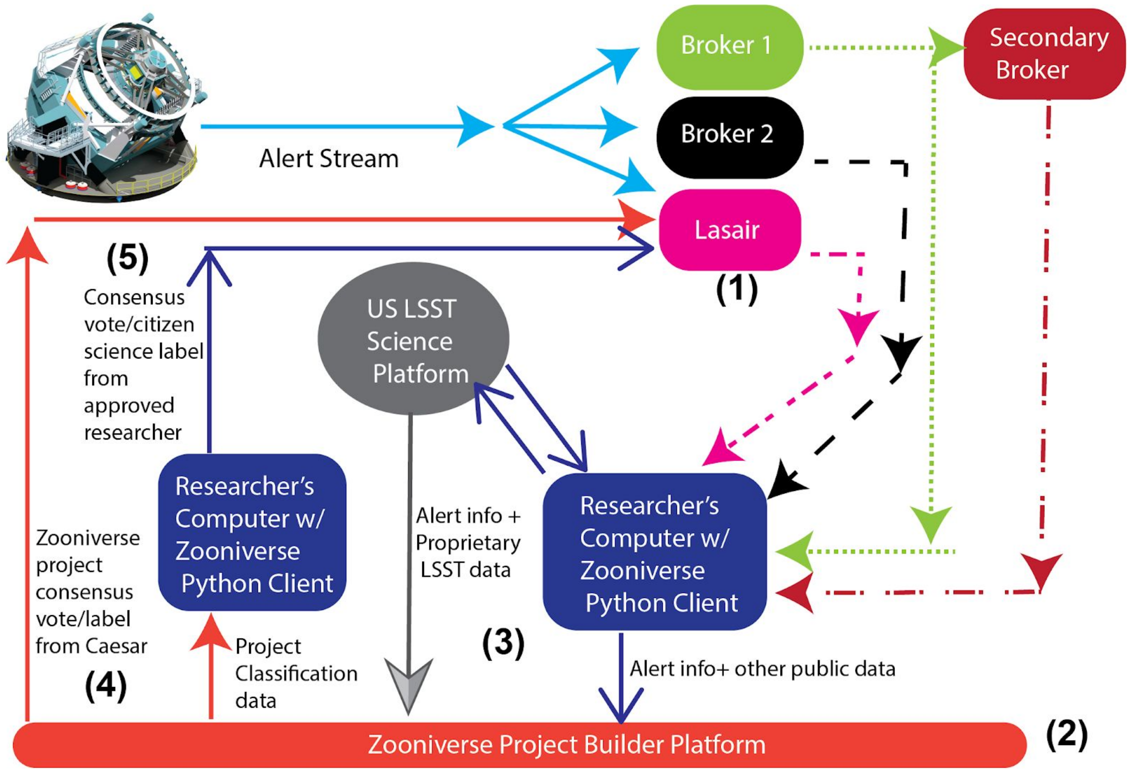

Figure 2: Lasair+Zooniverse diagram: Pathways for LSST alerts (or current ZTF alerts) to be filtered by Lasair and passed through a live stream to a researcher who then uploads the data to the Zooniverse platform and the researcher's Zooniverse citizen science project.

The detailed process is as follows:

1) The researcher creates an active query with Lasair that is then set up to output the filtered alerts via a Kafka stream. (Currently a researcher contacts the Lasair team to enable the Kafa stream functionality for ZTF Lasair.)

2) The researcher sets up a project and active workflow using the Zooniverse Project Builder. Zooniverse projects serve subjects and collects classifications from volunteers who perform assessment and classification tasks on the data served on the Zooniverse platform

3) The researcher uses the lasair_zooniverse python module[11] and the Zooniverse Panoptes python client[12] to populate the Zooniverse project with subjects derived from Lasair ZTF/LSST transient alerts to review. The lasair_zooniverse module has been developed by the University of Minnesota, Queen's University Belfast, and University of Oxford for a demonstration project on the Zooniverse platform. This open source code, documentation, and worked example is publicly available.

[11] https://github.com/rafia17/Zooniverse_SLSN
[12] https://github.com/zooniverse/panoptes-python-client

a) This module connects to and parses the Kafka stream produced by Lasair. Using the module, the researcher queries the standard Lasair interface, (currently via a html query) to obtain additional information about each alert, including light curve data and postage stamp images.

b) The alert data is converted using the python client to a format that can be ingested by the Zooniverse API[13] (Application Programming Interface) as subjects on the Zooniverse project.

c) We anticipate such code will be run regularly so that new alerts produced by ZTF/LSST can be uploaded on a nightly, weekly or less frequent basis. Decisions about what information to include in a Zooniverse subject will be made by the researcher running this code locally on the server/machine- some projects may want to show volunteers light-curves, or snapshot images, or both, and various decisions about if and when to display metadata also need to be made.

d) A researcher with LSST data access rights may at this stage choose to add elements from the Prompt Data Store or Data Releases (examples include context images or longer period light curves) if necessary. Thus, even if Lasair does not include proprietary LSST data, researchers can still add proprietary data to their Zooniverse project while using Lasair's advanced Sherlock filtering capabilities.

4) Subjects are classified by multiple volunteers on the Zooniverse project website. Once the classifications have met the relevant criteria, set by the researcher, the subject is retired, taken out of rotation on the Zooniverse website. The researcher either downloads the volunteer classification data through the Zooniverse Panoptes python client or uses the Caesar client[14], the Zooniverse's real time data processing platform, to combine the multiple volunteer assessments and decide on a label (e.g. supernovae or not a supernovae) for the subject image (Lasair alert).

5) For pre-approved Zooniverse projects: Using the Zooniverse API and a specific key, the researcher or the Caesar client posts the citizen science label to Lasair. Not all citizen science projects should post back to Lasair, but there may be select cases such as an LSST version of Supernova Hunters where the citizen science assessment would be useful to incorporate into Lasair.

The process described above is for the real time processing of LSST alerts, but this also works for 'Archive' projects using a historical record of the alert stream that can generate subject sets directly via a query to the Lasair service, without assembling; requests of order 10000 subjects are not expected to be constrained by the available resources.

Expanding the Zooniverse Python Client to Additional LSST Alert Broker Concepts

---

13 https://panoptes.docs.apiary.io/#

14 https://help.zooniverse.org/next-steps/caesar-realtime-data-processing/

We have developed the lasair_zooniverse python module to upload the contents from the Lasair kafa stream to the Zooniverse platform as a prototype for additional broker concepts serving ZTF data or future LSST data. We believe the scenario outlined above will make it possible for a researcher to develop citizen science projects using any approved LSST alert stream brokers' filtering and classification capabilities to select which LSST transients should have a crowdsourced review. Since the researcher uses a python module powered by the Zooniverse python client to parse and upload the realtime output from the brokers, there is no specific software architecture, API, or web service that needs to be added to the LSST brokers or the Zooniverse. It should be straightforward for additional LSST alert broker concepts (especially those currently using ZTF alerts as a testbed) to adapt the lasair_zooniverse module.The LSST broker teams would only need to develop and maintain a python module. This module could eventually be incorporated directly into the Zooniverse Panoptes python client. For further questions, we recommend contacting the Zooniverse team directly at contact@zooniverse.org.

References

Bellm E. ,Blum R., Graham M., et al. for the LSST Project. 2020, LDM-612; Plans and Policies for LSST AlertDistribution, Tech. rep. https://docushare.lsst.org/docushare/dsweb/Get/LDM-612

Ivezíc, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873,111
LSST Science Collaboration, Abell, P. A., Allison, J., et al.2009, ArXiv e-prints, arXiv:0912.0201

Kaiser N. et al., 2010, Proc. SPIE, 7733, 77330E

Narayan, G., Zaidi, T., Soraisam, M. D., et al. 2018, ApJS,236, 9

Nordin, J., Brinnel, V., van Santen, J., et al. 2019b, A&A, 631, A147Oke, J. B. 1974, ApJS, 27, 21

Patterson, M. T., Bellm, E. C., Rusholme, B., et al. 2019, PASP,131, 018001

Saha, A., Wang, Z., Matheson, T., et al. 2016, inProc. SPIE, Vol. 9910, Observatory Operations:Strategies, Processes, and Systems VI, 99100F

Smith K. W., Williams R. D., Young D. R. et al 2019 RNAAS 3 26

Wright, D. E., Lintott, C. J., Smartt, S. J., et al. 2017, MNRAS, 472, 1315