



**HAL**  
open science

## A data science platform to enable time-domain astronomy

Michael W Coughlin, Joshua S Bloom, Guy Nir, Sarah Antier, Theophile Jegou Du Laz, Stefan van der Walt, Arien Crellin-Quick, Thomas Culino, Dmitry A Duev, Daniel A Goldstein, et al.

► **To cite this version:**

Michael W Coughlin, Joshua S Bloom, Guy Nir, Sarah Antier, Theophile Jegou Du Laz, et al. A data science platform to enable time-domain astronomy. *Astrophys.J.Suppl.*, 2023, 267 (2), pp.31. 10.3847/1538-4365/acdee1 . hal-04094693

**HAL Id: hal-04094693**

**<https://hal.science/hal-04094693>**

Submitted on 9 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



# A Data Science Platform to Enable Time-domain Astronomy

Michael W. Coughlin<sup>1</sup> , Joshua S. Bloom<sup>2,3</sup> , Guy Nir<sup>2,3</sup> , Sarah Antier<sup>4</sup> , Theophile Jegou du Laz<sup>5</sup> ,  
 Stéfan van der Walt<sup>6</sup> , Ariën Crellin-Quick<sup>7</sup> , Thomas Culino<sup>8</sup> , Dmitry A. Duv<sup>7</sup> , Daniel A. Goldstein<sup>7</sup> ,  
 Brian F. Healy<sup>1</sup> , Viraj Karambelkar<sup>5</sup> , Jada Lilleboe<sup>1</sup> , Kyung Min Shin<sup>9</sup> , Leo P. Singer<sup>10</sup> , Tomás Ahumada<sup>5</sup> ,  
 Shreya Anand<sup>5</sup> , Eric C. Bellm<sup>11</sup> , Richard Dekany<sup>12</sup> , Matthew J. Graham<sup>5</sup> , Mansi M. Kasliwal<sup>5</sup> , Ivona Kostadinova<sup>5</sup> ,  
 R. Weizmann Kiendrebeogo<sup>1,4,13</sup> , Shrinivas R. Kulkarni<sup>14</sup> , Sydney Jenkins<sup>15</sup> , Natalie LeBaron<sup>2</sup> ,  
 Ashish A. Mahabal<sup>16,17</sup> , James D. Neill<sup>5</sup> , B. Parazin<sup>1,18</sup> , Julien Peloton<sup>19</sup> , Daniel A. Perley<sup>20</sup> , Reed Riddle<sup>12</sup> ,  
 Ben Rusholme<sup>21</sup> , Jakob van Santen<sup>22</sup> , Jesper Sollerman<sup>23</sup> , Robert Stein<sup>5</sup> , D. Turpin<sup>24</sup> , Avery Wold<sup>21</sup> , Carla Amat<sup>8</sup> ,  
 Adrien Bonnefon<sup>8</sup> , Adrien Bonnefoy<sup>8</sup> , Manon Flament<sup>8</sup> , Frank Kerkow<sup>1</sup> , Sulekha Kishore<sup>5</sup> , Shloke Jani<sup>1</sup> ,  
 Stephen K. Mahanty<sup>1</sup> , Céline Liu<sup>8</sup> , Laura Llinares<sup>8</sup> , Jolyane Makarison<sup>8</sup> , Alix Olliéric<sup>8</sup> , Inès Perez<sup>8</sup> ,  
 Lydie Pont<sup>8</sup> , and Vyom Sharma<sup>1</sup>

<sup>1</sup> School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455, USA; [cough052@umn.edu](mailto:cough052@umn.edu)

<sup>2</sup> Department of Astronomy, University of California, Berkeley, CA 94720, USA

<sup>3</sup> Lawrence Berkeley National Laboratory, 1 Cyclotron Road, MS 50B-4206, Berkeley, CA 94720, USA

<sup>4</sup> Artemis, Observatoire de la Côte d'Azur, Université Côte d'Azur, Boulevard de l'Observatoire, F-06304 Nice, France

<sup>5</sup> Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA

<sup>6</sup> Berkeley Institute for Data Science, University of California Berkeley, Berkeley, CA 94720, USA

<sup>7</sup> Weights and Biases, Inc., 1479 Folsom Street, San Francisco, CA 90063, USA

<sup>8</sup> ESILV - École Supérieure d'Ingénieurs Léonard de Vinci, Paris, France

<sup>9</sup> EnergyHub, Inc., 41 Flatbush Avenue, Suite 400A, Brooklyn, NY 11217, USA

<sup>10</sup> Astroparticle Physics Laboratory, NASA Goddard Space Flight Center, Code 661, Greenbelt, MD 20771, USA

<sup>11</sup> DIRAC Institute, Department of Astronomy, University of Washington, 3910 15th Avenue NE, Seattle, WA 98195, USA

<sup>12</sup> Caltech Optical Observatories, California Institute of Technology, Pasadena, CA 91125, USA

<sup>13</sup> Laboratoire de Physique et de Chimie de l'Environnement, Université Joseph KI-ZERBO, Ouagadougou, Burkina Faso

<sup>14</sup> Owens Valley Radio Observatory 249-17, California Institute of Technology, Pasadena, CA 91125, USA

<sup>15</sup> Department of Physics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

<sup>16</sup> Division of Physics, Mathematics and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA

<sup>17</sup> Center for Data Driven Discovery, California Institute of Technology, Pasadena, CA 91125, USA

<sup>18</sup> Northeastern University, Boston, MA 02115, USA

<sup>19</sup> IJCLab, Univ Paris-Saclay, CNRS/IN2P3, Orsay, France

<sup>20</sup> Astrophysics Research Institute, Liverpool John Moores University, IC2, Liverpool Science Park, 146 Brownlow Hill, Liverpool L3 5RF, UK

<sup>21</sup> IPAC, California Institute of Technology, 1200 E California Boulevard, Pasadena, CA 91125, USA

<sup>22</sup> Deutsches Elektronen-Synchrotron DESY, Platanenallee 6, D-15738 Zeuthen, Germany

<sup>23</sup> The Oskar Klein Centre, Department of Astronomy, Stockholm University, AlbaNova, SE-10691, Stockholm, Sweden

<sup>24</sup> Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, F-91191, Gif-sur-Yvette, France

Received 2023 April 28; revised 2023 May 30; accepted 2023 June 14; published 2023 July 31

## Abstract

SkyPortal is an open-source software package designed to discover interesting transients efficiently, manage follow-up, perform characterization, and visualize the results. By enabling fast access to archival and catalog data, crossmatching heterogeneous data streams, and the triggering and monitoring of on-demand observations for further characterization, a SkyPortal-based platform has been operating at scale for >2 yr for the Zwicky Transient Facility Phase II community, with hundreds of users, containing tens of millions of time-domain sources, interacting with dozens of telescopes, and enabling community reporting. While SkyPortal emphasizes rich user experiences across common front-end workflows, recognizing that scientific inquiry is increasingly performed programmatically, SkyPortal also surfaces an extensive and well-documented application programming interface system. From back-end and front-end software to data science analysis tools and visualization frameworks, the SkyPortal design emphasizes the reuse and leveraging of best-in-class approaches, with a strong extensibility ethos. For instance, SkyPortal now leverages ChatGPT large language models to generate and surface source-level human-readable summaries automatically. With the imminent restart of the next generation of gravitational-wave detectors, SkyPortal now also includes dedicated multimessenger features addressing the requirements of rapid multimessenger follow-up: multitelescope management, team/group organizing interfaces, and crossmatching of multimessenger data streams with time-domain optical surveys, with interfaces sufficiently intuitive for newcomers to the field. This paper focuses on the detailed implementations, capabilities, and early science results that establish SkyPortal as a community software package ready to take on the data science challenges and opportunities presented by this next chapter in the multimessenger era.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

*Unified Astronomy Thesaurus concepts: Open source software (1866); Astronomy software (1855); Time domain astronomy (2109)*

## 1. Introduction

The proliferation of data from synoptic sky surveys affords an ever-increasing discovery potential in time-domain astrophysics, on solar system, Galactic, and extragalactic scales. However, for this potential to be realized fully and efficiently, the volume, velocity, and heterogeneity of such data must be managed, curated, and coordinated. And, as often discovery requires more data for novel insights, new sky events demand an optimization of follow-up observations across a growing array of precious and loosely federated resources. We now recognize the need to change how astronomers analyze these data sets, with the application of data science principles to facilitate discovery and new insights.

Multimessenger astronomy, with the integration of data sets from, e.g., Advanced LIGO (Aasi et al. 2015) and Advanced Virgo (Acernese et al. 2015) for gravitational waves, e.g., IceCube (Aartsen et al. 2017) for neutrinos, e.g., the Zwicky Transient Facility (ZTF; Bellm et al. 2019; Graham et al. 2019; Masci et al. 2019; Dekany et al. 2020), and the forthcoming Vera C. Rubin Observatory (Ivezic et al. 2019) for optical astronomy, and  $\gamma$ -ray survey instruments such as the Neil Gehrels Swift Observatory mission (Gehrels et al. 2004) and Fermi’s Gamma-ray Burst Monitor (Fermi-GBM; Meegan et al. 2009), is unique and demanding for the number of integrated sources of data it requires to study targeted systems. As its quintessential example, the detection of GW170817 (Abbott et al. 2017a) and its associated electromagnetic transients AT2017gfo (Abbott et al. 2017b; Coulter et al. 2017; Smartt et al. 2017) and GRB 170817A (Abbott et al. 2017c; Goldstein et al. 2017; Savchenko et al. 2017) have demonstrated the power of multimessenger astronomy for measurements of the Hubble constant (Abbott et al. 2017; Hotokezaka et al. 2019; Coughlin et al. 2020a, 2020b; Dietrich et al. 2020), the neutron star equation of state, and  $r$ -process nucleosynthesis (Chornock et al. 2017; Coulter et al. 2017; Cowperthwaite et al. 2017; Pian et al. 2017; Rosswog et al. 2017; Smartt et al. 2017; Rosswog et al. 2018; Kasliwal et al. 2019; Watson et al. 2019) among many other science cases.

Previously, the large sky area of gravitational-wave localizations has made searching for (and doing science with) counterparts particularly challenging (e.g., Coughlin et al. 2019a; Andreoni et al. 2019; Goldstein et al. 2019; Gomez et al. 2019; Lundquist et al. 2019; Andreoni et al. 2020a; Ackley et al. 2020; Antier et al. 2020; Gompertz et al. 2020; Kasliwal et al. 2020; Anand et al. 2021). While localization accuracies should improve somewhat (Petrov et al. 2022) in the fourth LIGO–Virgo–KAGRA observing run (O4; Coughlin 2020), continued innovation in how searches are conducted remain paramount. To this end, we have undertaken an effort to integrate the capabilities from two “Marshals,” so named as software stacks designed to marshal candidates to follow-up telescopes, used during the third observing run (O3): the Global Relay of Observatories Watching Transients Happen (GROWTH) Target of Opportunity Marshal (Coughlin et al. 2019b) and the Global Rapid Advanced Network Devoted to the Multimessenger Addicts (GRANDMA) Interface and Communication for Addicts of the Rapid follow-up in the multimessenger Era (icare) platforms (Antier et al. 2019),

within SkyPortal, the open-source Target and Observation Manager (TOM) currently employed in ZTF Phase II (van der Walt et al. 2019). These TOMs and Marshals, whose role in the time-domain ecosystem will be discussed below, are designed to support the astronomical community to work efficiently with these data sets, identify sources of interest, and obtain follow-up. In the following, we will use “TOM” to describe systems of this type, as opposed to “Marshal,” which can refer to systems more narrowly focused on follow-up, but acknowledge that these two terms are nearly synonymous in the community.

In this paper, we describe the components built within SkyPortal, focusing on their application to multimessenger astronomy. There are two existing platforms which deploy SkyPortal that heavily rely on these features: (i) *fritz*,<sup>25</sup> the platform used by the ZTF collaboration for Phase II operations and (ii) *icare*,<sup>26</sup> the platform used by the GRANDMA collaboration. These are private instances deploying SkyPortal for use by members of their collaborations; we will point out aspects that are deployment specific below to give readers a sense of what aspects may be deployment specific. We describe our perspective on the time-domain ecosystem and how SkyPortal fits in it in Section 2. The key features and technical implementation of the software stack are discussed in Section 3. Descriptions of the *fritz* and *icare* deployments, as well as an example analysis applied to ongoing  $\gamma$ -ray burst (GRB) searches are shown in Section 4.2. We summarize our conclusions and future outlook, including our perspectives on software development in an academic environment, in Section 5.

## 2. The Time-domain Software Ecosystem

We briefly review the time-domain software ecosystem here, and encourage interested readers to explore the references to understand this continually evolving ecosystem. Understanding this ecosystem is required for understanding the role of SkyPortal; our vision for the role of this software stack in this ecosystem is presented by Figure 1. The flowchart indicates the interrelated chains of research infrastructure, both hardware and software, within which SkyPortal functions.

There are a number of other related software stacks available with some overlap with SkyPortal’s capabilities. For example, YSE-PZ is a platform with data-broker-like querying and filtering abilities, and that offers a turn-key solution to teams looking for a focus on transient survey management (Coulter et al. 2023). The TOM Toolkit focuses on customization for different science use cases, deployable both as a standalone package as well as with its functionality integrated into larger ecosystems (Street et al. 2018). Astro-COLIBRI is an application programming interface (API)-based platform, accessible via a public website or through mobile apps, which provides a user-friendly overview of transient astronomical events (Reichherzer et al. 2021). Astro-COLIBRI enables users to identify possible multi-wavelength or multimessenger counterparts quickly, and assess potential ground-based follow-up options.

<sup>25</sup> <https://github.com/fritz-marshall/fritz>

<sup>26</sup> <https://github.com/grandma-collaboration/icare>

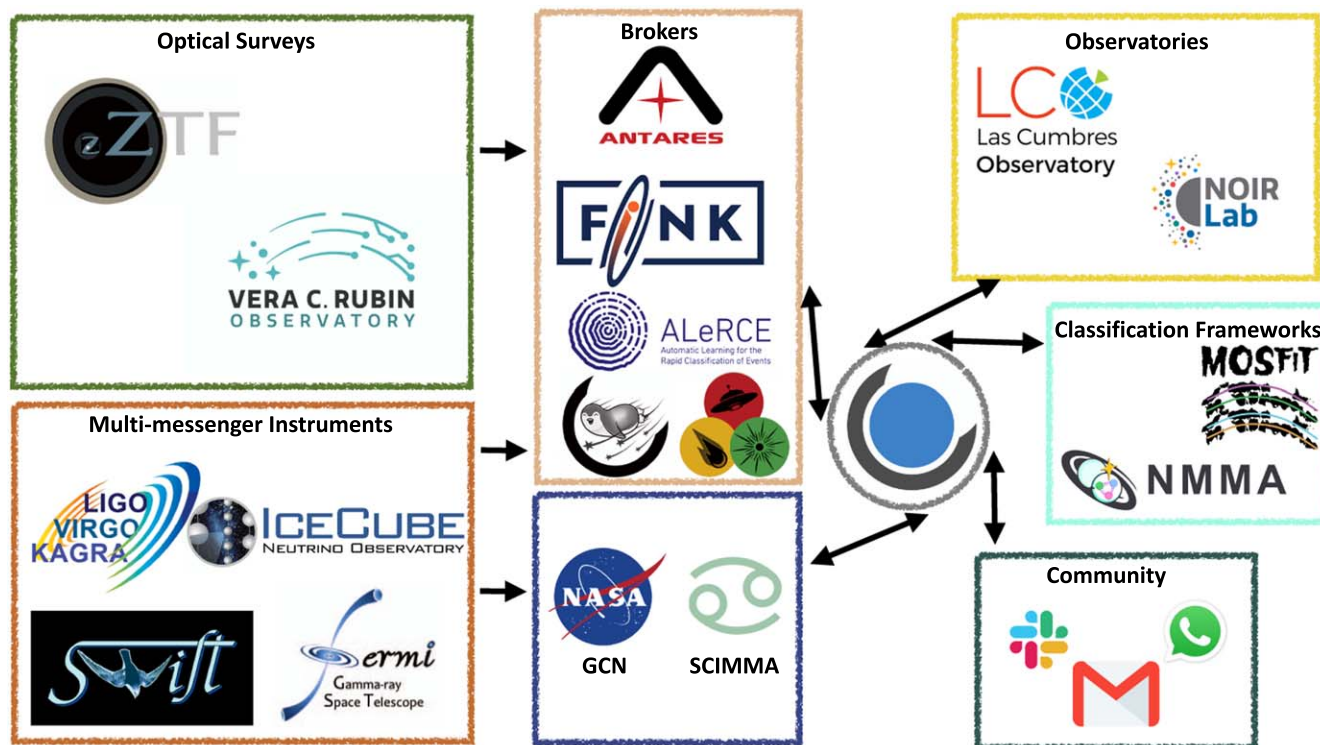


Figure 1. Flowchart for the multimessenger, technical ecosystem envisioned in this platform.

In some ways, SkyPortal has ambitions to be a sophisticated “full-stack transient ecosystem” that goes beyond a TOM of these kinds—integrating capabilities of telescope scheduling, observing optimization, spatial catalogs, astronomical data analysis, and others (see Section 3) into a single software stack.

The first key piece of the software ecosystem are brokers like Fink (Möller et al. 2020) or ALerCE (Förster et al. 2021), which ingest alerts from optical surveys such as ZTF (Bellm et al. 2019; Graham et al. 2019) or the Rubin Observatory (Ivezić et al. 2019) and provide value-added annotations such as crossmatching with other catalogs. A purpose of a broker is to act as a layer between alert issuers and the scientific community analyzing the alert data. It exposes one or more services to help the scientists to (manually and programmatically) analyze the alert data from telescopes and surveys efficiently. Generally, brokers collect and store alert data, enrich them with information from other surveys and catalogs or user-defined added values such as machine-learning classification scores, and redistribute the most promising events for further analyses, including follow-up observations. Large-scale surveys such as ZTF and the Rubin Observatory have required broker teams to design and implement new technological approaches to operate in real time on large computing infrastructures for a wide variety of science cases. Recently, the Rubin Observatory announced which community brokers will have unrestricted access to the complete alert stream for the next decade:<sup>27</sup> ALerCE (Förster et al. 2021), AMPEL (Nordin et al. 2019), ANTARES (Matheson et al. 2021), BABAMUL and Fink (Möller et al. 2020), Lasair (Smith et al. 2019), and Pitt-Google. Most of these brokers are already operating on the ZTF alert stream, which has

constituted a prime opportunity to engage projects with the scientific community while preparing the ground for the upcoming Rubin Observatory alert data.

Some of these alert brokers also process alert data from other surveys and telescopes, including data at different wavelengths and from different messengers, and interested readers are encouraged to explore the references above for specifics. These annotations may include information from other wavelengths, spectroscopy, or historical data; they also often provide classification information, such as using machine-learning classifiers that account for these annotations. All of these innovations are designed to support astronomers in purifying the alert streams to restrict the objects to those of scientific interest for particular users.

However, the role of a broker is not to ingest all the data that a user of such a broker might need for a complete analysis, as it is sometimes beyond the expertise or the road map of the broker team, or it would require a substantial amount of work to integrate, or simply because some data require proprietary access. Instead, broker systems expose interoperable tools so that users can export enriched data elsewhere to continue the analysis. In practice, this is where TOMs and Marshals, for which interfaces with most brokers have been developed, come in, where the coordination of follow-up observations and further scientific analyses will take place after the brokers enrich the data.

The second key piece of the software ecosystem relates to platforms whereby multimessenger instruments (e.g., high energy, gravitational waves, neutrinos, and transient phenomena in general) can share alerts and rapid communications; these currently include the General Coordinates Network (GCN; Singer & Racusin 2023) and the Scalable Cyberinfrastructure to support Multimessenger Astrophysics<sup>28</sup> (SCiMMA) project.

<sup>27</sup> <https://www.lsst.org/scientists/alert-brokers>

<sup>28</sup> <https://scimma.org/>

These projects serve to ingest and then redistribute these alerts to subscribing astronomers in the community. In this sense they are similar to brokers, although in general do not seek to annotate the ingested data streams, as the brokers do; it is again TOMs and Marshals that are often ingesting these streams, although some brokers also enable alert stream crossmatching with these multimessenger notices.

TOM systems interact with both brokers and these multimessenger alert distribution systems. While there is inevitably some redundancy with brokers especially, the focus of TOMs is to be a centralized database whereby coordination of follow-up observations of identified candidates (such as those passing the filters run within brokers) and support of astronomer science (such as integrating classification frameworks to be used on the resulting data sets) is possible. These TOM systems involve both human and algorithmic interactions, such as for the prioritization and submission of follow-up requests for target lists, as well as software analysis interfaces to light curve fitting codes like MOSFiT (Guillochon et al. 2018) or the “Nuclear Physics—Multimessenger Astronomy” (NMMA) framework (Pang et al. 2022).

TOMs provide an important conduit for information flow within this ecosystem. While many of these services are developed independently, they rely on well-defined interfaces such as APIs to interact. For time-domain astronomy in particular, they must also interact with flexible algorithmic frameworks which can plot and fit transients such as supernovae or tidal disruption events, as well as periodic and quasi-periodic objects such as eclipsing binaries. They also need to interact with core external services to different astronomical communities such as the Transient Name Server<sup>29</sup> (TNS) or the Minor Planet Center<sup>30</sup> (MPC). While it is possible to downselect the services a TOM interacts with, it is inevitable that these communities overlap, given that, for example, transient astronomers are always interested in avoiding following up asteroids like those reported to the MPC.

Underlying the software components of the time-domain ecosystem are the methods by which data are stored, normalized, and exchanged. While image and tabular data are well-served by international (IVOA<sup>31</sup>) standards for accessing, querying, and exchange, the only aspect of the time-domain ecosystem that is currently described is the reporting of astronomical transients (VOEvents; Seaman et al. 2011). There is as yet no agreed data model, query protocol, or exchange format for time series<sup>32</sup> or the semantically rich heterogeneous data items of annotations (human comments, classifications, etc.). The semistructured nature of the latter of these means as well that they are difficult to map into existing regular data standards. SkyPortal manages heterogeneous input by enforcing data exchange with a strict API schema which can be accessed by any programmatic tool (and could be described as a VOSI interface for compliance; Graham et al. 2017). However, there is clearly a strong need within the burgeoning multimessenger astronomy (MMA) community for a reengagement in defining schema, formats, and services for multimessenger data exchange.

<sup>29</sup> <https://www.wis-tns.org/>

<sup>30</sup> <https://minorplanetcenter.net/>

<sup>31</sup> <https://www.ivoa.net>

<sup>32</sup> A time-series exchange protocol, VOTimeseries, was proposed (Bloom et al. 2008) but never adopted by the IVOA.

### 3. SkyPortal Framework

In the following subsections, we present the key features and specific, implemented subsystems of SkyPortal in use to enable applications in time-domain astronomy. These features are available in every SkyPortal instance, upon download and deployment of the software stack.

#### 3.1. Associated Software Packages

By design, SkyPortal is open source, and through its development, changes have been contributed to open-source, upstream packages such as `sncosmo` (Barbary et al. 2016); while it would be out-of-scope to describe all of the packages on which SkyPortal relies, we briefly describe two such packages, `baselayer` and Hierarchical Equal-Area isoLatitude Pixelization-Alchemy (HEALPix-Alchemy; Singer et al. 2022), for which development is particularly coordinated.

##### 3.1.1. *Baselayer*

The open-source package `baselayer`<sup>33</sup> provides the backbone of a generic science-application web framework on top of which SkyPortal is built. Originally built for the open-source time-domain machine-learning application `cesium` (Naul et al. 2016), `baselayer` implements, e.g., websocket communication between the back end and front end, authentication, access control, microservices, configuration, process management, cron jobs, migrations, and external logging. As far as possible, it combines standard solutions, such as Tornado, Nginx, supervisor, and Python Social Auth, instead of implementing them from scratch. By delegating web infrastructure to `baselayer`, concerns are better separated and the SkyPortal team can focus on domain-specific concerns. That said, the SkyPortal team also contributes platform improvements upstream as required, exposing these to a wider audience.

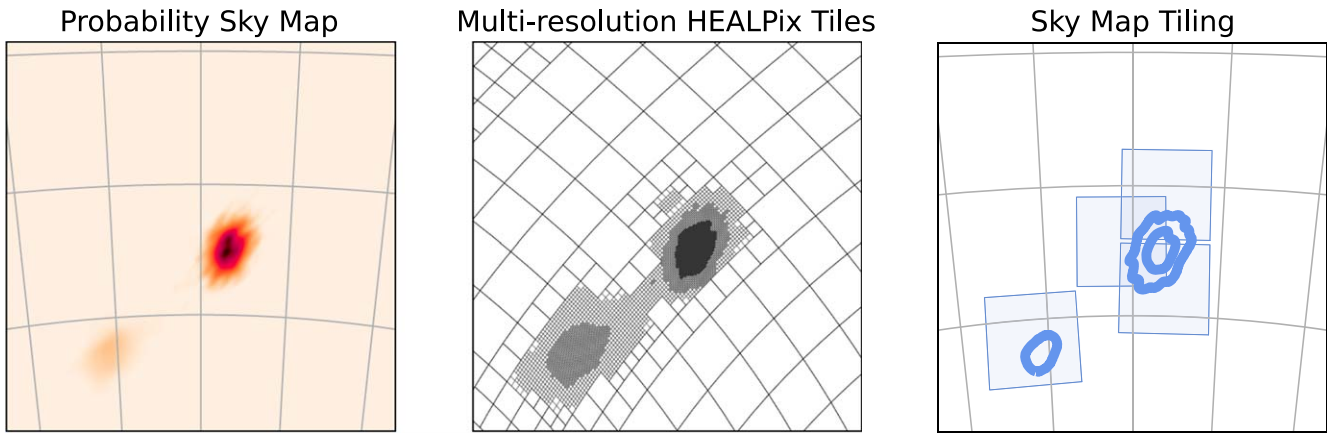
At the core of `baselayer`, we find `sqlalchemy`,<sup>34</sup> an object relational mapper written in Python, allows easy interaction with a relational database, here being PostgreSQL<sup>35</sup> (Stonebraker & Rowe 1986). Several base models—which are the abstractions of tables in the database—are defined to introduce the concept of users, permissions, roles, API tokens, and groups. An add on of Python Social Auth provides the necessary models for user authentication using many `openid` authentication services. In both `baselayer` and SkyPortal, authentication through Google services is the method of choice. `sqlalchemy` also allows handling multiple parallel and separate connections to the database named sessions. This is an important requirement for any API-based application, where multiple users need to read/write/modify/delete from the database simultaneously. Wherever interacting with the database is necessary, we open a session, add/modify/delete from that session, and then commit the changes. Changes are made to the database only once the commit method is called, and PostgreSQL is capable of resolving conflicts between multiple commit sessions if necessary.

For the users to interact with the API-based application that is `baselayer`, we define API handlers, which are simply

<sup>33</sup> <https://github.com/cesium-ml/baselayer>

<sup>34</sup> <https://www.sqlalchemy.org/>

<sup>35</sup> <https://www.postgresql.org/>



**Figure 2.** An example of the multi-resolution HEALPix sampling scheme for tiling localizations. The left panel shows an example heat map image representing the localization probabilities; darker, deeper colors represent higher probability densities. The middle panel shows the boundaries of the multi-resolution HEALPix tiles on which the sky map was sampled. The right panel shows potential telescope field of views covering that sky map with the sky map confidence level contours included.

Python classes with the standard GET, POST, DELETE, PUT, and PATCH methods an API is expected to have. Then, Tornado maps each handler to the associated API endpoint. It is capable of calling the right API method of the handler and passing in-path parameters for each API call. In SkyPortal, every time a new feature is needed, database models and handlers that interact with them can be added to the application following the same process. When it comes to permissions, baselayer provides Python decorators that can be specified above each API method to specify what permissions or roles a user needs to use it. Similarly, database-level access restrictions can be defined in every sqlalchemy model, to ensure that users only have access to the data points that have been shared with them or with the groups to which they belong.

Using a microservice architecture managed with supervisor, we can run multiple instances of the application in parallel to ensure availability and reduce downtime. That way, the app can easily scale horizontally based on the number of users or when new features are added. Moreover, we can add microservices to run background operations continuously. In baselayer, this is used to run the application of course, but also the database migration manager, the webpack builder, websocket server, cron jobs, external logging, and nginx. In SkyPortal, this has been proven to be extremely useful when adding computationally expensive or long-running features such as the ingestion of GCN events with low latency, processing of observation plans, sending notifications and reminders, and programming recurrent API calls. These services will be discussed in further detail later in this paper.

### 3.1.2. HEALPix-Alchemy

To ensure calculations are sufficiently computationally expedient in SkyPortal, we represent the complex geometry of regions of interest (such as multimessenger sky maps) and the instruments' fields of view using the HEALPix (Górski et al. 2005) framework; to this end, we have developed a PostgreSQL extension enabling HEALPix-based cross-matches to make rapid processing on these large alert databases possible (Singer et al. 2022), named HEALPix-Alchemy. We briefly describe the innovation and its use within SkyPortal here, and encourage readers of this paper to see Singer et al. (2022) for further details.

In version 14, PostgreSQL introduced a new multirange type, consisting of an array of ranges, with a fast aggregation function, which takes ranges as its input and returns their union as a multirange; when combined with the multiorder-coverage (MOC) representation of sky maps (Fernique et al. 2014), this tandem yields highly efficient database-side processing of sky map overlaps.

Most multimessenger events do not provide us with a precise localization to observe, but instead a map of probability density (see, e.g., Figure 2) using MOC to divide the sky into tiles of diverse resolutions. Each tile has an associated probability density that the event occurred in it. MOC uses the HEALPix tessellation algorithm to divide the spherical observable sky into 12 diamond shapes of equal area. Each of them can be recursively divided into four to obtain the level of resolution needed to represent the variation of probability from one point of the sky to another; recursively dividing these tiles into four subdiamonds increases the order and therefore the resolution.

At order 0, tiles have indexes from 1 to 12, and at order 1, from 1 to 64. For instance, tile 1 at order 0 is now made of tiles 1–4 at order 1, and tiles 1–16 at order 2, and so on recursively until the highest order is order 29, where a tile side size is only 0.38 milliarcseconds; this minimum resolution is set by requiring that the corresponding pixel index can be stored as a 64 bit signed integer without overflow. Thanks to the recursive indexing schemes of the HEALPix algorithm, when saving MOC maps in the database, we can represent each tile as the range of tiles that it is composed of at order 29. This is convenient as it allows us to represent tiles as a probability density associated with a range set. To do so, HEALPix-Alchemy combines the power of HEALPix and SQLAlchemy, a Python package used to communicate with PostgreSQL databases using Python. HEALPix-Alchemy uses the range sets feature of PostgreSQL databases to represent tiles and single points with 64 bit integers and provides the tools necessary to crossmatch tiles and points. For software stacks that rely on relational databases, this range sets representation of HEALPix tiles is the best (and only viable) implementation to perform rapid spatial queries.

As MOC has become the standard representation for sky maps in multimessenger astronomy (Fernique et al. 2022), HEALPix-Alchemy enables efficient queries of instrument footprints for rapid crossmatches with historical GRB, neutrino, and gravitational-wave sky maps. Similarly, based

on the coordinates of any object, observation, or galaxy, `SkyPortal` calculates and stores its associated `HEALPix` index at order 29 (which is the highest order, as stated above). This means that the spatial crossmatch between sources, galaxies, and multimessenger events is performed simply by looping over the localization tiles of an event localization to perform trivial intersection operations. Furthermore, thanks to the range set representation of `PostgreSQL`, performing the intersection and union of the `HEALPix` tiles representing regions of space is equivalent to a simple merging of sorted lists of integers (Singer et al. 2022). These intersection operations indicate if a given `HEALPix` index (a point) is contained in any of the range sets of the indexes corresponding to `HEALPix` tiles (a spatial region).

Using the same mechanism as the sky maps, `SkyPortal` also represents instrument fields using MOC representations of their footprint on the sky. The same basic union operations can be performed to determine which portions of a sky map (or in principle, objects and galaxies) an instrument is able to observe. Because the `HEALPix` algorithm has limits in resolution, the usage of MOC to define circles, polygons, and other objects with a basic geometry is in some sense an approximation (Fernique et al. 2014); nevertheless, the MOC standard is an efficient way of representing instrument fields in a compact and generic manner.

### 3.2. Point-source Features

While the line between features designed to interact with individual point sources and those designed to describe and interact with larger areas of the sky is blurry, we still find it useful as a broad diagnostic for understanding subsets of the features. Here, we describe the workflow for studying individual sources, from filtering to photometry to follow-up.

#### 3.2.1. Filters and Scanning

With a full night of observing, ZTF sends  $\sim 600,000$ – $1,200,000$  alerts per night, with over 2,000,000 alerts sometimes possible (Patterson et al. 2018); for comparison, the Rubin Observatory expects to send  $\sim 10,000,000$  alerts per night. Given that the overwhelming majority of these alerts are not of interest to the average user, a combination of automated filtering and human scanning is used to reduce the alert list to a short list of sources deemed worthy of further attention.

Within `SkyPortal`, every science objective has a dedicated “group” of users, and one associated filter per group. These filters are defined within the brokers described in Section 4.1. As brokers are filtering the alert stream automatically in real time, sources producing alerts that pass a given filter are known as candidates. `SkyPortal` can be configured to save immediately all candidates that pass the filter to a group (a collection of sources), but in most cases, users will prefer to vet the stream manually to separate sources of genuine interest from false positives in a process referred to as candidate scanning. All candidates passing a filter during a chosen period (e.g., the past 24 hr or the past three nights) can be displayed rapidly on the Candidates page. This page displays image cutouts for each candidate location from ZTF and other surveys, along with light curves and some additional information (coordinates, TNS crossmatches, etc.) and links to other resources. The user can peruse this list of candidates

and save only those of interest based on this additional information. Candidates that are not saved will reappear if the source continues to produce alerts that pass the filter in the future, but the user can also choose to hide any sources that are clearly false positives from their future scanning efforts.

#### 3.2.2. Photometry, Spectroscopy, and Photometric Series

In addition to the ingestion of photometry from alert streams provided by brokers, `SkyPortal` benefits from both public photometry services from Pan-STARRS Data Release II (Chambers et al. 2016; Flewelling 2018) and the forced photometry services of both ZTF (Masci et al. 2019) and ATLAS (Tonry et al. 2018; Smith et al. 2020). These are available for querying through their API services. The visualization interface for photometry has been described previously (van der Walt et al. 2019), but briefly, it is a `Bokeh`<sup>36</sup> implementation that allows one to view either magnitude or flux as a function of time, with tool tips providing basic information about the individual photometry points. There is also a panel that allows the user to specify a potential period to phase-fold the photometry, in the case of periodic sources.

Statistics on each source are collected and updated upon the insertion of new data points. This includes, e.g., the brightest magnitude, the time when the source was last detected, and so on. These statistics are updated on insertion to avoid the costly recalculation over hundreds of data points per source. Such statistics are used to filter sources quickly based on their light curves, without needing to calculate each statistic for each source in each query.

Another way to store large photometric data sets has recently been added to `SkyPortal`: the photometric series. This data product will be used by surveys with very high cadences and continuous coverage of the same field, e.g., the Transiting Exoplanet Survey Satellite (Ricker et al. 2014) takes images at a maximum cadence of 20 s over a 27 day series. In this case, saving each of the  $\sim 10^5$  data points as individual rows in the database becomes unmanageable, both in terms of storage and in terms of communicating the data in and out of `SkyPortal`. Instead, we store the underlying data as HDF5 files and keep a queryable record in the main database of the file (along with some statistics).

Spectra are uploaded either manually or via API. The spectra are associated with each source but the platform also enables the users to provide comments, allowing users to discuss either the source or the individual spectra. There are also standard line lists available to perform line identification, enabling classification.

#### 3.2.3. Finding Charts and Starlists

To aid the rapid follow-up of sources of interest, `SkyPortal` enables photometric and spectroscopic finding charts to be generated on the fly. In interactive mode, a user selects a catalog image from DESI DR8, ZRF Reference Images, DSS2, or PS1, as well as the image size and number of spectroscopic offset stars. Using a flux-weighted median position from all photometric detections, up to four nearby isolated stars from Gaia DR3 are selected as spectroscopic offset stars. The offsets are calculated using the parallax and proper motion of the stars

<sup>36</sup> <https://docs.bokeh.org/en/latest/>

from Gaia DR3 assuming the current datetime as the observing epoch. These offset stars can be copied from the front end in different starlist formats. Finding charts can be generated programmatically and from the front end (single click) with default parameters. A downloadable PDF is assembled and cached to facilitate fast observing run organization.

### 3.2.4. Periodograms

For variable stars, *SkyPortal* enables an interactive periodogram analysis. An adapted JavaScript version of the generalized Lomb–Scargle implementation (Zechmeister & Kürster 2009) allows for client-side computation of the periodogram. A user can change photometry filters and interactively select trial periods as well as visualize the folded light curve on the selection period and period aliases.

### 3.2.5. Source Classifications

*SkyPortal* offers a classification system that lets users tag sources with relevant labels. The namespace for classifications is set by taxonomies, which are collections of labels connected with each other using the nested JSON format. *SkyPortal* sets a default site-wide taxonomy for use in all groups on the application based on the *tdtax*<sup>37</sup> repository. Users can also post custom taxonomies (that adhere to a strict taxonomy schema) to specific groups, limiting their classifications' visibility to members of those groups. Each classification has an associated probability that quantifies confidence in the label.

Based on the system built within the ZTF variable star Marshal (Coughlin et al. 2021; van Roestel et al. 2021), *SkyPortal* allows users to post classifications from a drop-down menu one by one or use a slider interface that facilitates faster labeling. The slider interface initially shows the top-level classifications of the selected taxonomy. Each classification is accompanied by a slider to adjust its probability. If the slider of a classification is moved to a nonzero probability, the child classifications of the parent will be shown along with additional sliders. This interface allows multiple classifications to be posted at once.

*SkyPortal* also features a classification voting system that facilitates collaborative labeling of sources in a prompt but informative way. Users can express their confidence in each existing classification visible to them by casting a thumbs “up” or “down” vote. When a user adds or deletes a source’s classification or casts a vote, *SkyPortal* marks the source as “Labeled.” The user can later choose to view only unlabeled sources in a group to allow quick resumption of the labeling process.

### 3.2.6. Follow-up

From within *SkyPortal*, API-based triggering is available for the telescope of the Las Cumbres Observatory network (Street et al. 2018), the Katzman Automatic Imaging Telescope (Li et al. 2003), the Liverpool Telescope (Steele et al. 2004), Swift (Gehrels et al. 2004), the SED Machine (SEDM; Blagorodnova et al. 2018; Rigault et al. 2019) on the Palomar 60 inch telescope, and the Wide-field Infrared Transient Explorer (WINTER; Lourie et al. 2020). Adding a new telescope’s follow-up capability is straightforward. Figure 3 shows an example drop-down menu customized for SEDM.

Groups that scan for candidates within *SkyPortal* often use the same platform to trigger photometric and spectroscopic

The image shows a web form for scheduling follow-up observations. At the top, it says "Follow-up" with an upward arrow. Below that is an "Allocation" dropdown menu showing "Palomar 1.5m / SEDM - Program A (PI Michael Coughlin)". Underneath is a "Share Data With" section with two radio buttons, "Program A" and "Program B", where "Program A" is selected. The "Mode" dropdown is set to "4-shot (ugri)". The "Priority" field contains the number "1". The "Start Date (UT)" field is "06/18/2022" and the "End Date (UT)" field is "06/25/2022". At the bottom of the form is a blue "SUBMIT" button.

Figure 3. JSON-based follow-up interface.

follow-up in the same portal. Priorities, time windows, and observational setups can be selected and edited. Once the data have been obtained, some systems such as the robotic SED Machine (Blagorodnova et al. 2018) automatically reduce (Rigault et al. 2019) and upload the spectra for visualization; as a deployment-specific example, in the case of *fritz*, if a source spectrum is automatically matched to a thermonuclear spectrum (supernova Ia; Fremling et al. 2021), those spectra and classifications are also directly uploaded to the TNS as part of the Bright Transient Survey (Fremling et al. 2020; Perley et al. 2020). For other sources, human interaction allows for the classification (or rescheduling) of other targets, as well as the inclusion of detailed comments and discussions to share results and collaborate on particular objects.

## 3.3. Non-Point Source and Multimessenger Features

While many of the features described above have applicability to sources of all types, including infrastructure for follow-up, some of the features have been developed focusing on interactions with larger areas of the sky, as appropriate for many multimessenger science cases. Here, we describe those features, focusing on their applicability for supporting multimessenger source analysis and follow-up.

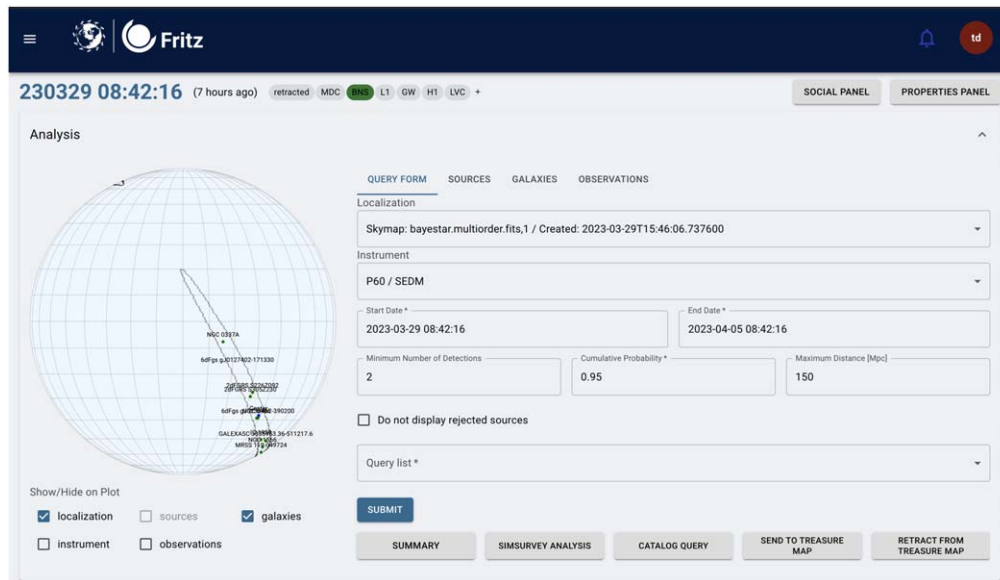
### 3.3.1. GCN Events

Much of the existing infrastructure for communicating multimessenger observations originated from the GRB detection community in the 1990s (Barthelmy et al. 1995), with the decades-old Gamma-Ray Coordinates Network continuing to play a central role in modern time-domain astronomy. This system includes the machine-readable GCN Notices, which uses three legacy communication formats (text, binary, and VOEvents; Allan et al. 2017), each with its own communication protocols. Recently, there have been concerted efforts to reimagine this system as a modern “General Coordinates Network,” transitioning to a unified communication protocol using open-source Apache Kafka, with the ultimate aim of relaying all data via the industry-standard JSON format.<sup>38</sup>

<sup>37</sup> <https://github.com/profsb/timedomain-taxonomy>

<sup>38</sup> <https://gcn.nasa.gov/>





**Figure 4.** GCN Page—Analysis Section: here, users can visualize the localizations associated with an event, and query the sources, galaxies, instrument tiles (for a given allocation on an instrument), and observations contained in the different localizations, while restraining their searches based on several parameters: cumulative probability, number of detections, distance in megaparsecs, and first and last detection dates. Other features accessible in this section allow the user to query catalogs of candidates from different alert brokers, run simulations with `simsurvey`, and generate text-based summaries similar to GCN circulars with added version control. It is also possible to submit the executed observations to `treasuremap.space` (Wyatt et al. 2020). Side panels can be opened that contain the properties of the event and localizations, light curves, GCN notices and circulars, as well as a comment section for users to communicate their results.

SkyPortal is fully integrated with this new GCN Kafka format and can ingest notices shared via this protocol. Which of the available streams should be ingested is specified in the configuration of the application. These GCN notices typically contain either full MOC sky maps (Fernique et al. 2022), coordinates in R.A. and decl. along with an associated error radius, or the vertices of a region shaped like a polygon. In all cases, a new entry is added to the database containing the localization itself (transformed to an MOC for efficient and standardized queries) and other available event properties such as localization area, median distance, or astrophysical probability. These events then can be subsequently queried, crossmatched to sources or galaxies, and used to coordinate subsequent follow-up.

SkyPortal supports the analysis of these events by allowing for different queries within the localization map: telescope observation footprints, galaxies, and transient sources. The analysis page of each event can display the error regions in a spherical projection, and can additionally display the results of the queries (see an example in Figure 4). Moreover, the sources returned in the query will appear listed in the Sources table of the event page, enabling further visual inspection. The live interaction that occurs in the Sources menu has shown to be essential to enable candidate vetting and follow-up prioritization. For example, by clicking the arrow next to the name of the source or candidate, SkyPortal displays the available photometry, as well as cutouts of the coordinates of the source. Once a candidate association has been confirmed or ruled out, its status can be updated and comments can be left for future reference or internal communication among collaborators. Finally, SkyPortal can use this information to generate automated GCN circulars which summarize observations performed by telescopes that overlap the event, the associated sources detected by those observations, and the estimated event localization probability covered by those observations. Figure 4 shows the interface associated with the GCN analyses.

### 3.3.2. Observations: Planning and Execution

A central component of multimessenger astronomy is the efficient coordination of follow-up for a given event, especially for gravitational-wave and GRB triggers which are initially poorly localized. Astronomers may have access to both wide field-of-view instruments and narrow field-of-view instruments of different depths and wavelength coverages. Developing strategies for optimal follow-up with a particular telescope has been an area of active research in recent years (e.g., Coughlin et al. 2018, 2019c; Almualla et al. 2020).

SkyPortal has extensive functionality to perform these optimizations automatically in a user-friendly GUI. Upon ingestion of localizations, SkyPortal can generate optimized observing schedules using the open-source Gravitational-wave Electromagnetic Optimization (`gwemopt`) software package (Coughlin et al. 2018, 2019c). `gwemopt` can generate custom plans for individual telescopes, or broader network plans which coordinate multiple telescopes simultaneously. It can either generate an agnostic sky-tiling schedule, or use internal catalogs to generate a galaxy-targeting schedule while accounting for user-tunable factors such as airmass, start and end time, or available filters.

Plans are typically generated by default at the time of event ingestion, with the option for further plan generation via a configurable web form. Each of these plans can be visualized on the event page, overlaid with the initial localization. Observation plans can be triggered automatically, but can also be triggered manually from the GCN event page.

To assess the efficacy of the generated plans, we use the open-source simulator software `simsurvey`<sup>39</sup> (Feindt et al. 2019), and are working with the author of `skysurvey` to integrate it.<sup>40</sup> In particular, we use `simsurvey` to simulate

<sup>39</sup> <https://github.com/ZwickyTransientFacility/simsurvey>

<sup>40</sup> <https://github.com/MickaelRigault/skysurvey>

**Figure 5.** GCN Page—Observation Plans: here, users select an allocation, the fields of the instrument of the allocation that overlap with a localization, as well as pick the parameters to use from the MMA API class of the instrument. Multiple observation plans can be created simultaneously and combined if necessary. Moreover, buttons allow users to download observability and airmass charts, as well as recompute airmass in real time to show which tiles of the sky map are observable. It is also possible to submit the proposed observations to `treasuremap.space` (Wyatt et al. 2020). Once created, all observation plans will be listed by instrument, where they can be updated, sent, or deleted to/from an instrument queue.

transients of various types (kilonovae, GRB afterglows, supernovae, etc.) to estimate the probability of detecting one of these transient types within the sky map, given the proposed observation plan. For kilonovae, the optical/near-infrared counterparts to binary neutron star mergers generated from the radioactive decay or  $r$ -process elements (Metzger 2017), we use a POSSIS-based (Bulla 2019) grid of kilonova models spanning the plausible parameter space for kilonovae from binary neutron star and neutron star–black hole mergers (Dietrich et al. 2020; Bulla 2023). For GRB afterglows, we use *afterglowpy* (Ryan et al. 2020), an open-source computational tool modeling forward shock synchrotron emission from relativistic blast waves as a function of jet structure and viewing angle. This analysis accounts for properties of the sky maps (i.e., probability in R.A. and decl., and distance in the case of gravitational-wave sky maps), as well as properties of the observations (i.e., R.A. and decl. for each field, observation time, limiting magnitude, filters, etc.).

To send back data from the observations based on these plans, telescope teams can use the API of *SkyPortal*, or add them directly on the Observation Page. Observation plans will automatically be crossmatched with the newly added observations, so they can appear on the list of observation plans for the event. If all observations scheduled in a plan are finished, the observations will be labeled as completed. Due to the elusive (red and fast-fading) nature of gravitational-wave and GRB counterparts, several consecutive follow-up campaigns could result in a limited number of counterpart detections. Thus, one of the main science cases in multi-messenger astronomy is using upper limits to constrain

transient properties (Coughlin et al. 2019d, 2020c). To quantify our sensitivity to electromagnetic counterparts from our past observations, we again use *simsurvey*. We inject light curves consistent with the distance and location information contained in the sky maps and calculate the detection efficiency (i.e., the ratio between the number of detected transients and the number of injected transients) based on the field-by-field limiting magnitudes of our uploaded observations. We can then report these detection efficiencies, along with the median depth of our observations promptly via GCN. Figure 5 shows the interface associated with the observation planning.

### 3.3.3. Galaxy Catalogs

Due to the relatively local nature of multimessenger astronomy, both follow-up focusing on galaxy catalogs (Arcavi et al. 2017; Valenti et al. 2017) and comparison of transient locations to catalogs is important. Within the database, galaxies can have several attributes: name, alternative name, catalog name, R.A., decl., redshift ( $z$ ), and distance in megaparsecs (if known), as well as other properties and their associated margins of error if provided. Concerning the usage of catalogs, as deployment-specific examples, in *fritz*, we use the Census of the Local Universe (CLU; Cook et al. 2017) catalog, which is complete to 85% in star formation and 70% in stellar mass at 200 Mpc; this catalog currently remains proprietary. In *icare*, GLADE+ (Dalya et al. 2021), an open-source catalog containing more than 22 million galaxies, complete up to 44 Mpc and nearly 90% complete at even 500 Mpc, is used. The latter is a combination of six separate astronomical catalogs: GWGC, 2MPZ, the Two Micron All

Sky Survey (2MASS) XSC, HyperLEDA, and WISExSCOSPZ.

While we focus on these two catalogs, the API interface makes ingestion of other galaxy catalogs (and therefore crossmatching) straightforward, with a minimal requirement of R.A., decl., and galaxy name. As one can imagine, the large size of these catalogs drives the need to perform efficient crossmatches between objects, multimessenger event localizations, and these galaxies; our innovations with HEALPix-Alchemy has helped, however, these are challenges that we are still working on to date, with performance improvements in the road map for O4. However, initial versions suitable for production performance are already implemented and functional.

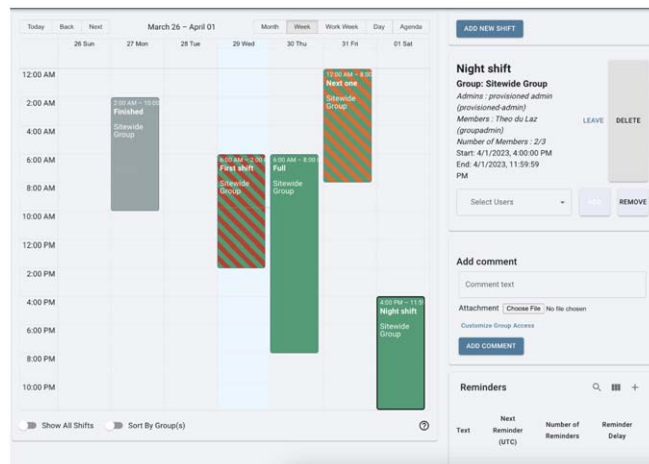
### 3.4. Interpretation and Coordination Features

There are some features in the application with utility irrespective of the source(s) of interest. These include person-power coordination, notification, and analyses. Here we describe the SkyPortal infrastructure built to address such components of the MMA workflow.

#### 3.4.1. Shift Scheduling

The scheduling of person-power is essential to implementing a robust multimessenger or time-domain astronomy campaign; time is of the essence when it comes to triggering the follow-up of a source or planning the observations necessary to cover the localization of an event. Within SkyPortal, we have introduced the concept of “shifts,” whereby each day is divided into shorter periods—usually in four periods of 6 hr—where teams of follow-up advocates and “shiffters,” whose time zone best accommodates such a period, are assigned.

A page dedicated to the shifts has been added to SkyPortal to allow the creation, visualization, and management of shifts through a Google Calendar-like page. On the shift calendar, one can create a shift, join or leave a shift, delete a shift, comment on a shift, and add or remove people (for those who have the admin role). Each shift has an associated group and max shifter capacity. If said capacity is not reached within 48 hr before the start of the shift, this will be indicated on the calendar, and members of the group will be notified as well. Moreover, our experience during the previous O3 campaign showed us that the planning of a shifter could change at the last minute, preventing them from covering their shift and breaking the continuous monitoring of new events. To solve this problem, a feature allowing a shifter to request a replacement from other members of their group by notifying them was added. When selecting a shift on the calendar, one can see the shift management panel, the comments panel, the reminders panel, as well as the shift summary panel at the bottom of the page. On the shift summary element, a list of events that occurred during the selected shift—if any—will appear. By clicking on an event, one can see the list of sources contained within the latest localization of the event. The sources listed are the ones that have been first and last observed within one week after the first detection of the event. When leaving a comment associated with a shift in its comment section, one can mention/tag events, sources, and other users. This feature allows us to centralize and gather the work of shifters on SkyPortal in real time while associating their messages with objects and events. This capability has replaced



**Figure 6.** Shifts Page: users can see the shifts on a calendar and click on them to display more information and interact with them, as well as create new shifts.

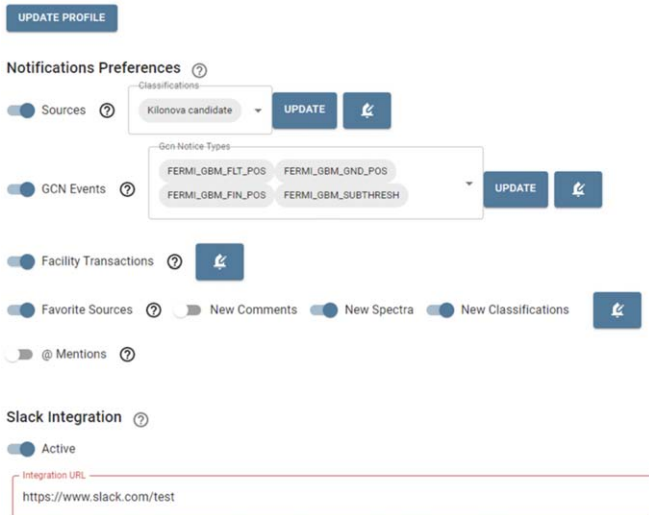
some need for collaborations to rely on external tools such as Slack, which has been the option of choice so far for this type of scheduling process. Figure 6 shows the interface associated with the shift planning.

#### 3.4.2. Notification Framework

As previously mentioned, we rely on notifications to keep users informed about recent activities within the app in a user-tunable fashion, given that not all users need to be notified of the same activities. To accomplish user-defined notifications, we built a notification framework in SkyPortal, where each user can specify their needs on the user profile page. On this page, one can choose to be notified of the following:

1. When a given classification is added to any source, based on a list of classifications in which the user is interested.
2. When sources tagged as “favorites” record different activities selected by the user, such as new comments, new spectra, and new classifications.
3. When a new GCN event or new GCN notice for an event already in the platform is ingested. The user can select which GCN notice types they want to be notified for, as well as specify properties of either the event itself (e.g., classification as a binary neutron star) or the sky map (e.g., size of the localization region).
4. When a new facility transaction is added, including follow-up requests and observation plans.
5. When a user is mentioned by another user in a comment.

For each of these notification types, in addition to the in-app notifications, the user can choose to be notified by email, Slack, SMS, WhatsApp, and/or phone call. By email and Slack, a user will always be notified. When the SMS, Whatsapp, and/or phone call option is selected, users must choose if they want to be notified while on shift, and/or every day but during a specific time period (i.e., from 8 AM to 10 PM local time). For instance, during a recent observation campaign, shifters activated notifications for new events of the notice types that the campaign was interested in (in that case, Swift notifications specifically). Moreover, the possibility of “bot” notifications to create alert channels on Slack was added; these work through the user providing a slack channel URL to interact with as a webhook. We also automatically subscribe some users to



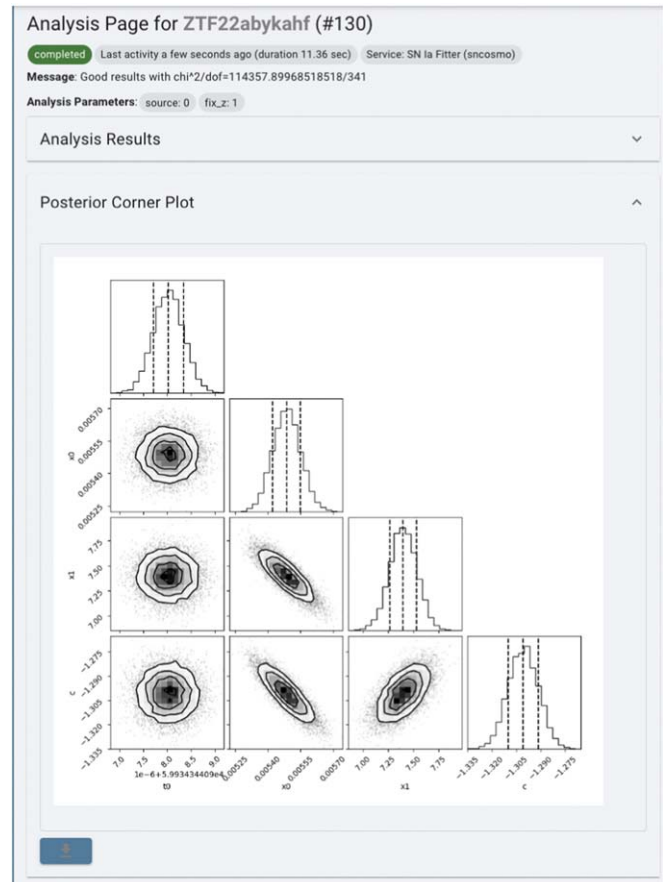
**Figure 7.** User Profile–Notifications Preferences: users can activate or deactivate notifications, specify parameters on which to notify, as well as choose where to be notified by clicking the bell icon next to each notification type.

receive certain critical notifications; for instance, SkyPortal always notifies the admins of an allocation when a new facility transaction has been created on the instrument for which they are admins. This is done to ensure that new follow-up requests or observation plan submissions cannot be missed by those responsible for triggering the associated observations on their telescopes. Figure 7 shows the interface associated with the notification framework.

### 3.4.3. Analysis Platform

A common way to characterize transients in optical astronomy is through the analysis of photometric, multiband light curves. A variety of Bayesian inference frameworks have sprung up to meet the needs of transient characterization, including MOSFiT (Guillochon et al. 2018), the NMMA framework (Pang et al. 2022), among others. These frameworks are used to calculate Bayes’ factors for the purpose of model comparison as well as posteriors for transient characterization. While time-of-discovery analyses can reside within the purview of brokers, analysis of follow-up photometry and spectra can provide critical insights.

SkyPortal now enables a wide range of analyses on individual sources, viewing analysis packages as standalone web services capable of receiving source data and returning analysis summaries (in JSON format), parameter posteriors (in `xarray arviz.InferenceData` format; Kumar et al. 2019; Figure 8), and/or diagnostic plots. Site administrators first create analysis services, by providing a URL (and authentication credentials) to an existing service, the schema of required and optional parameters, and data types to be sent from SkyPortal to the analysis service. When an analysis is initiated (either via the front end or via API), all data viewable by the user in the analysis data types are packaged and sent to the analysis service. Via a unique unauthenticated webhook, the analysis service then asynchronously returns analysis results. (Analysis results generated without SkyPortal-packaged source data and without a webhook can also be saved by an authenticated token.) The results are stored in a persistent datastore by SkyPortal and can be retrieved by



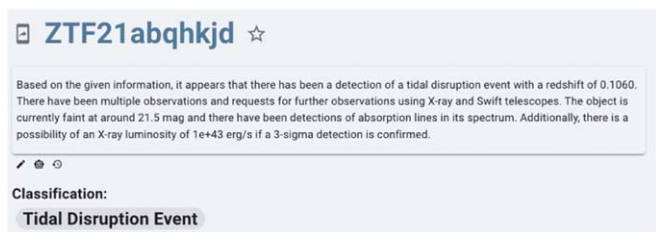
**Figure 8.** Analysis posterior for a fit to ZTF 22abykahf with the Type Ia supernova analysis service.

analysis ID or by querying for analysis results by source name. Posterior corner plots and other returned plots are also rendered on individual analysis pages.

While the analysis platform enables remote third party services, SkyPortal also ships with prebuilt analysis services that are run locally as microservices. A supernova service wraps the `sncosmo` (Barbary et al. 2016) fitter, and an afterglow/kilonova service wraps the `nmma` (Pang et al. 2022) fitter.

### 3.4.4. AI Summarization and Embeddings

One challenge for end-users in the face of hundreds (or thousands) of active events is maintaining a current understanding of the state of knowledge and activity for each event. To facilitate “quicklook” insights into sources, following the emergence and general availability of highly capable large language models (LLMs), the SkyPortal team released a prebuilt analysis service to provide human-readable summaries of individual sources, using redshift, classifications, and comments. When such an analysis is initiated on a source, the summarization service formats the available source data along with a predefined prompt requesting a summary; the OpenAI API (`openai.ChatCompletion`) returns a summary which is then displayed on the source (Figure 9). This summary can be easily edited, with the summary history being saved. A sitewide API key to OpenAI can be used or individual users can configure their own key and specific completion parameters. SkyPortal defaults to use the `gpt-3.5-`



**Figure 9.** Source page visualization of the AI-generated summary of using the ChatGPT 3.5 completion service on ZTF 21abqhkjd (Yao et al. 2021).

turbo model but users can opt to employ the more capable (but currently slower) gpt-4 model.

Once a summary is created, an OpenAI embeddings service (using the `text-embedding-ada-002` model) is called with the summary that returns a vector embedding of the source in  $n = 1536$  dimensions. These embeddings are saved in a `pinecone` database.<sup>41</sup> Using these embeddings, similar sources can be retrieved by rank-ordering cosine distance metrics or natural language queries (first embedded, then compared against the summary corpus) can be answered. Future embeddings will incorporate photometry, spectra, and annotations to facilitate further AI-guided exploration of `SkyPortal` data sets.

#### 4. Deployments, Science Validation, and First Results

In this section, we describe the specifics of two `SkyPortal` implementations, `fritz` and `icare`, as well as a worked example of an “offline” search for GRB afterglows from Fermi-GBM with ZTF. `SkyPortal` implementations mainly differ on (i) where they are deployed and (ii) the data stream that they receive, and so in these cases, the brokers that they rely upon to filter and post results from ZTF. For (i), `fritz` is deployed on Google Cloud Platform, while `icare` runs on dedicated hardware on site at IJCLab in Orsay, France. For (ii), we describe these in the following subsection.

##### 4.1. Alert Brokering

Several `SkyPortal` instances rely on alert brokers to populate their databases with candidates that users can scan and save as sources (or reject) for their groups. In the case of `fritz`, it uses `Kowalski`,<sup>42</sup> an open-source, multisurvey data archive and alert broker (Duev et al. 2019) for performing filtering and crossmatches, including rejection of bogus objects (Duev et al. 2019); some science groups also rely on `AMPEL` (Nordin et al. 2019). In the case of `icare`, it relies on `ZTF-Fink` (Aivazyan et al. 2022), including its filtering capabilities targeting fast transients, to select promising candidates.

##### 4.1.1. Kowalski

`Kowalski` comprises a nonrelational (NoSQL) `MongoDB`<sup>43</sup> database to store ZTF alerts and other catalogs, and an API layer that allows external users (such as `fritz` itself, or the users using it) to interact with it. Among other benefits, `MongoDB` provides fast execution times ( $\sim \log(N)$ ) for standard operations, allows for efficient positional queries, and saves individual entries in the binary JSON format—well-

matched with the AVRO format of ZTF alerts. A dedicated Kafka consumer in `Kowalski` consumes the ZTF alert stream at IPAC in real time and ingests alerts into the database, where they are stored as “documents.” `Kowalski` provides an infrastructure relying on parallel computing to filter and enhance many alerts at once prior to ingestion. Basic filters can be specified in the configuration, on top of which user-based filters can be defined for the different instruments handled by `Kowalski`. Those can be sent to `Kowalski` using its API, allowing the development of GUIs (like the dedicated filter page on `Skyportal`) where users can define filters in a human-readable JSON format (see Section 3.2.1).

Users define filters in the `MongoDB Query Language (MQL)` syntax, so that they can be run on the alerts stored in `Kowalski` as “documents.” A filter takes the form as a `MongoDB` “aggregation pipeline” that is tuned to a specific scientific objective. An aggregation pipeline is a list of dictionaries, each representing a specific stage in the filtering process. A stage performs an operation on the document, such as calculating a value (e.g., the age or color of the transient) or filtering the document (e.g., only pass documents with  $\text{color} > 1$  mag.). All documents are fed to the first stage, but only the output of a stage is fed to its subsequent stage—this ensures that only documents of interest are retained by the end allowing for efficient automated filtering of alerts.

As part of the filtering document, users can specify which alert fields can be posted to `SkyPortal` as annotations, allowing for quick analysis of source properties. For alert enhancement, `Kowalski` can be configured to crossmatch alerts with other catalogs and candidates from the other instruments or run machine-learning inference on them. The user-defined filters are applied after the crossmatches and machine-learning scores are added, meaning that filtering can be done based on those values and not only the original alert fields. Based on the filtering results, an alert sentinel will post candidates and annotations. Once the enhanced alerts are ingested into `Kowalski`’s database, users can query them by running MQL-based queries using `Kowalski`’s API. The package `penquins`<sup>44</sup> provides a Python client to interact with the `Kowalski` API. `Kowalski` is containerized using `Docker` allowing for simple deployment on the cloud or on premises.

For `fritz`, `Kowalski` serves as an “upstream” aggregation pipeline that combines all photometric history, crossmatch data, and machine-learning scores for a ZTF alert available into a single document is run and then passed to the user-defined filter. Filters saved on `fritz` are version controlled and can be viewed/updated on a dedicated tab.

The instance of `Kowalski` that serves as the back end for `fritz` is deployed on premises at Caltech. As of 2023 March, it stores  $\sim 38$  TB of data including  $\sim 555$  million ZTF alerts and  $\sim 20$  other catalogs (e.g., 2MASS; Cutri et al. 2003; and ALLWISE; Cutri 2012). It processes millions of requests daily, received either from `fritz` or  $\sim 10$  individual users for whom direct access is allowed. It handles candidate filtering for  $\sim 80$  ZTF experiments (e.g., the Bright Transient Survey; Fremling et al. 2020; Perley et al. 2020) by running filters saved on `fritz` (see Section 3.2.1) on all new ZTF alerts in real time. In addition to ZTF, this instance also manages the alert stream for Palomar Gattini-IR (De et al. 2020)—an infrared time-

<sup>41</sup> <https://www.pinecone.io/>

<sup>42</sup> <https://github.com/skyportal/Kowalski>

<sup>43</sup> <https://www.mongodb.com/>

<sup>44</sup> <https://github.com/dmitryduev/penquins>

**Table 1**  
Kowalski

Parameter	Value
Data Stored	~38 TB
ZTF Alerts	$\sim 5.55 \times 10^8$
Other Catalogs	~20

**Note.** Key numbers for the Fritz instance of Kowalski.

domain survey—in a similar manner and may do the same for WINTER (Lourie et al. 2020). These figures are summarized in Tables 1 and 2.

#### 4.1.2. AMPEL

In addition to Kowalski, some science working groups in ZTF use AMPEL (Nordin et al. 2019) to process the alert stream and display the results on Fritz.

AMPEL divides its processing into four tiers. Each tier is handled by a set of *units* (pluggable, configurable modules) that each perform a specific task. At Tier 0 (*ingest*), alert packets are filtered, and selected alerts are decomposed into constituent data points (e.g., individual photometric points) and stored. At Tier 1 (*combine*), data points are combined into *states* (e.g., light curves). At Tier 2 (*augment*), data points and states are enhanced with auxiliary or derived information, e.g., catalog matches for photometric points or model fits for light curves. At Tier 3 (*react*), documents are aggregated by *stock* (transient object being tracked) and can be reported to an external service, e.g., to request follow-up for the  $N$  most interesting objects, or simply provide a target list for a traditional scanning workflow. Users can customize each of these tiers to create a *channel*, a real-time analysis plan specific to their intended science program. This could consist of, e.g., a custom alert filter, data-point association scheme, set of light curve analyses, and reporting scheme. Tiers 0–2 are driven by the alert stream, while Tier 3 is run at fixed intervals.

AMPEL integrates with SkyPortal via SkyPortalPublisher, a Tier 3 unit distributed as part of the Ampel-ZTF package.<sup>45</sup> Each stock that SkyPortalPublisher processes is saved as a candidate for a set of filters. It assumes a one-to-one mapping between AMPEL channels and (dummy) filters stored on Fritz by default, but a custom set of filters may be configured instead. Additionally, a stock is saved as a source in a set of target groups if so configured. Light-curve analysis results from Tier 2 are serialized to JSON and posted as comments on the source. This behavior predates the analysis platform described in Section 3.4.3, and we expect to migrate the analysis reporting of AMPEL to this service in the near future. If photometry and image cutouts were loaded as part of the Tier 3 process, these are posted as well. This makes it possible to use AMPEL with a standalone SkyPortal instance, but this is disabled in the ZTF instance, as competing back ends tend to confuse users and lead to unnecessarily high database loads. Finally, SkyPortalPublisher keeps a record of its communication with SkyPortal in the journal associated with each stock, and only attempts to post if it was updated since the last successful attempt.

Some science programs require two-way communication, e.g., the ZTF Nuclear Transients group uses classifications that

**Table 2**  
fritz

Parameter	Value
Photometry	~400 million
Candidates	~7.3 million
Sources	~420,000
Spectra	~11,000
Groups	~500
Users	~320
Tokens	~160
Filters	~80
Telescopes	~70
Instruments	~85
Comments	~170,000
Annotations	~3.2 million
Thumbnails	~12.2 million
GCN events	~3100

**Note.** Key numbers for the Fritz instance of SkyPortal.

human astronomers post to Fritz in their AMPEL-based scanning workflow. AMPEL enables this via e.g., FritzReport, a Tier 3 complement stage that injects a source record from SkyPortal into stock records before they are provided to downstream units.

#### 4.1.3. Fink

The GRANDMA collaboration uses the Fink broker (Möller et al. 2020) to select promising alerts from the public ZTF alert data stream and display the results on icare.

The current Fink platform works in four steps. First, alerts from multiple streams are continuously ingested and stored on disk. Second, alerts satisfying the quality cuts defined by the broker team are processed by a set of science modules. These science modules—currently a dozen, spanning solar system objects to Galactic and extragalactic science—are independent processing units developed by the community of users, and deployed in the Fink platform. They can work on a single input alert stream, or combine several streams together. These science modules enrich the initial alert packets using several techniques such as crossmatch with external catalogs of astronomical objects, or classification using machine- or deep-learning-based algorithms. All added values are made public. Third, enriched alert packets are filtered based on their content, and the most promising events are redistributed to the scientific community in real time. The filtering is again community driven, and users design and deploy filters to receive tailored information in real time. Finally, all enriched alert packets are stored in a database for permanent access and for further analyses. Fink exposes a number of services to the users to design modules, filters, and access the data.<sup>46</sup>

Fink integrates with icare with the SkyPortal Fink Client.<sup>47</sup> Alerts are first polled using the Fink Livestream service, and then pushed to icare using core functionalities of SkyPortal. The SkyPortal Fink Client runs as an optional microservice inside icare which simplifies the deployment and the maintenance for users. We expect to migrate this optional microservice to the main SkyPortal

<sup>45</sup> <https://github.com/AmpelAstro/Ampel-ZTF>

<sup>46</sup> <https://fink-broker.readthedocs.io>

<sup>47</sup> <https://skyportal-fink-client.readthedocs.io>

codebase in the near future so that it can benefit the SkyPortal community at large.

In the context of GRANDMA, three specific Fink filters targeting alerts associated with fast transients have been deployed (Aivazyan et al. 2022). One of the filters is based on the output score of a dedicated science module implementing a fast transient classification algorithm, and all filters were used in two subsequent follow-up campaigns of optical transients to search for orphan kilonovae (Biswas et al. 2022). Selected alerts are automatically published into *icare*, where continuous human scanning takes place (see below for the scanning capabilities within the platform). Astronomers would then decide to trigger further observations based on the received alerts’ content, but also using external information collected by the GRANDMA collaboration and ingested into *icare* directly by other means.

#### 4.2. Example Science Results

As mentioned above, as a first demonstration of the utility of the platform, we perform an “offline” search for GRB afterglows from Fermi-GBM with ZTF using *fritz*. ZTF has employed an ongoing program to search for GRBs (Coughlin et al. 2019b; Ahumada et al. 2021, 2022) and gravitational waves (Coughlin et al. 2019a; Anand et al. 2021) using “triggered” Target of Opportunity (ToO) observations, which use timing and/or localization information from other wavelengths or messengers. It is similarly possible to use “serendipitous” observations akin to those performed within routine survey observations, which have neither localization nor explosion time information, to find counterparts as well (Andreoni et al. 2020b; Ho et al. 2020; Andreoni et al. 2021).

We use the short duration GRB 200826A (Ahumada et al. 2021) to recreate the offline analysis of afterglows. For this, we download the Fermi XML VOEvent files and post them to *fritz*. Later, we trigger a customized query to post candidates within the days following an event to *fritz*. The spatial crossmatch employed within PostgreSQL (Section 3.1.2) is used to keep candidates within the localization region. Limiting our queries to extragalactic fast transients, we keep transients at high Galactic latitude ( $|b_{\text{Gal}}| > 10$  deg) with a fading rate faster than  $0.3 \text{ mag day}^{-1}$  in at least one band (Andreoni et al. 2021). The results of our query are displayed in the localization map, showing ZTF20abwysq among the sources recovered.

The first real-time ToO search using *fritz* was in 2022 November. We observed the localization region of the long GRB 221110A (trigger 689784662) detected by the GBM on the Fermi satellite with ZTF. We obtained a series of *g*- and *r*-band images covering 470 square degrees beginning at 09:47 UT on 2022 November 11 ( $\sim 15$  hr after the burst trigger time). This corresponded to  $\sim 62\%$  of the probability enclosed in the Earth-occultation-corrected GRB localization map. Each exposure was 240 s, reaching *g*- and *r*-band median depths of 21.2 mag and 21.1 mag, respectively. The images were processed in real time through the ZTF reduction and image subtraction pipelines at IPAC (Masci et al. 2019).

We queried the ZTF alert stream using Kowalski (Duev et al. 2021) and AMPEL (Nordin et al. 2019). We required at least two detections separated by at least 15 minutes to select against moving objects. The candidates within the 95% probability contour of the GRB localization map discovered

for the first time after the GRB trigger time, and which have more than two detections are displayed in Figure 10.

Further analysis consists of crossmatching our candidates with the MPC to flag known asteroids, reject stellar sources (Tachibana & Miller 2018), and apply machine-learning algorithms (Mahabal et al. 2019). These selection criteria as well as human vetting left us only with one candidate: ZTF22abtfyhi/AT2022aaaa. This transient was discovered at  $g = 20.0$  mag, but later found to be rising (and therefore uninteresting) after SEDM optical follow-up.

Reperforming this check across a variety of serendipitous observations of GRB sky maps yields the candidate quantities shown in Figure 11.

## 5. Development Ethos and Conclusions

In this paper, we have described the current state of multimessenger functionalities within SkyPortal and potential future functionality. The rapid approach of O4 will provide an opportunity to identify additional needs for enabling these tools to allow near-automated identification, follow-up, and characterization of potential counterparts. As an easily extensible framework, we plan a variety of improvements to the application; for example, we plan to be the first adopter of M4OPT, the Multi-Mission Multimessenger Observation Planning Toolkit package,<sup>48</sup> which is inspired by a previously prototyped Integer Linear Programming (ILP)-based schedule for ZTF (Parazin et al. 2022).

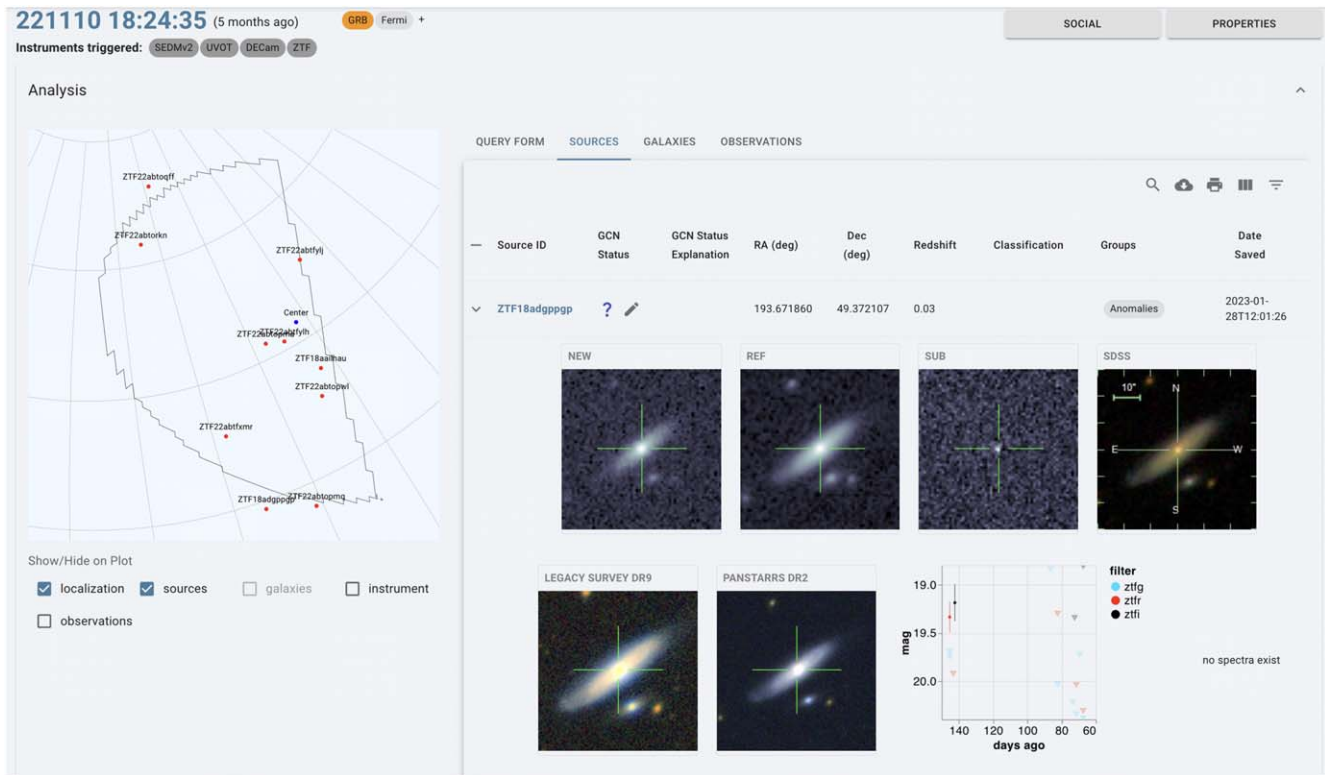
We have described herein a number of front-end-focused workflows, recognizing that science in the MMA era, while greatly enhanced by software and analysis tools, remains fundamentally a collective human endeavor. Recognizing that many workflows also happen programmatically outside of SkyPortal, the major functionality also is accessible via a well-documented API infrastructure.<sup>49</sup> It will come as no surprise that working across teams in different institutions—sometimes with aligned and sometimes with competing incentives—requires modern frameworks like SkyPortal to enable the establishment and curation of strict data access controls. Every photometry, spectroscopy, comment, annotation, and classification entry in SkyPortal is permissioned by group.

SkyPortal is, first and foremost, scientific software developed to enable astronomical data science. However, a close secondary goal is to train students in proper software development practices, which has been identified as a gap between coursework and eventual readiness for academic and industry practice (Licorish et al. 2022). These required skills include software engineering skills (e.g., identifying requirements, prototyping, designing, and using version control) as well as transferable and soft skills (e.g., teamwork and communication).

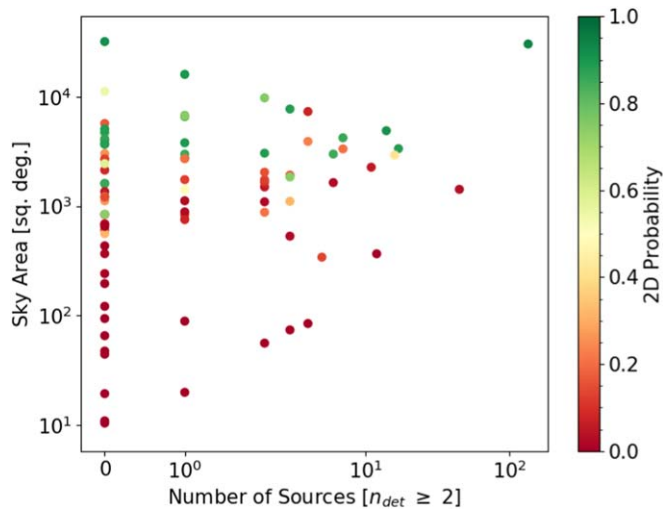
SkyPortal has had significant contributions from students at a variety of levels, with the faculty/permanent researchers associated with the project focused on bringing students into the fold. For example, two sets of 10 students from the Leonard de Vinci Engineering School have participated in the development of features, focusing on the *icare* version of the application. This particular effort has yielded multiple internships at member institutions for these students.

<sup>48</sup> <https://github.com/m4opt/m4opt>

<sup>49</sup> <https://skyportal.io/docs/api.html>



**Figure 10.** Analysis section from the GCN event page of GRB 221110A. The GBM region is uploaded after correcting for Earth occultation, which produces a cut on the shape of the original map. The sources detected in the region covered by ZTF are shown as red dots, with their names. The Sources menu allows for in-depth scanning of the candidates, showing cutouts of ZTF, and other surveys, as well as the photometry of the candidate. For this event, all the sources were unrelated to the GRB.



**Figure 11.** Scatter plot of sky area in square degrees vs. number of candidates (with at least two detections). The colors correspond to probability.

However, bringing on students continues to raise a number of challenges for a small project like SkyPortal. First of all, the onboarding efforts for an application the size of SkyPortal can be challenging. For this reason, the documentation is focused on both starting the application, making basic changes, and then an example of creating a pull request is quite thorough and continues to be updated as students join the project. However, significant dedicated senior developer time is still required not only initially during the onboarding process as well as later on as the students have become comfortable

contributing; domain knowledge and project scope, as well as addressing user requests, still require significant input. This is potentially problematic given that these academic projects are led by professors with many other demands on their time.

Additionally, the international nature of this project puts an added strain on these dynamics. In addition to natural challenges with language barriers, finding times when the entire team can meet is very challenging, and in practice, multiple meetings per week with limited overlaps in persons due to time zone differences creates some challenges with ensuring a cohesive project direction.

This is further challenged by active interactions with users, which occurs in a variety of venues; as for any project, features and bug fixes are done in conjunction with user requests, which are important for sustaining both the health of the application (for reporting bugs that are not caught in the unit testing) as well as needed features. These user interactions occur in a variety of ways, including through a mix of Announcements through GitHub discussions, e-mail updates, Google forms, Slack and other means of communication. This level of flexibility is convenient for users, but can be challenging for developers to track feedback in a variety of forms, not only the form the communication takes but also the level of feedback (i.e., small bug fixes versus large feature requests).

In this way, SkyPortal can serve as an important scientific platform for this new era of multimessenger astronomy. We expect that software platforms of this type will become prevalent to meet the needs of the scientific user community. We expect many projects will face similar challenges as their user communities grow and students contribute; we hope that the introduction of these projects for



students and others will continue to serve as important opportunities to do science while also learning best practices in a technical environment.

















### Acknowledgments

The authors appreciate comments from Dave Coulter, Fabian Schussler, and the anonymous referee on an initial draft of this paper.

M.W.C., J.L., and V.S. are supported by the National Science Foundation with grant No. OAC-2117997; M.W.C. is also supported by PHY-2010970. M.W.C. and R.W.K. were supported by the Preparing for Astrophysics with LSST Program, funded by the Heising–Simons Foundation through grant 2021-2975, and administered by Las Cumbres Observatory. The Gordon and Betty Moore Foundation, through both the Data-driven Investigator Program and a dedicated grant to develop SkyPortal, provided critical funding for this project without which this project could not have succeeded. J.S.B., G. N., A.C.-Q., and D.A.D. were partially supported by the Gordon and Betty Moore Foundation. B.P. thanks the Northeastern Lawrence Co-op Fellowship for their continuous support. J.P. thanks LSST-France and CNRS/IN2P3 for supporting Fink. The Kilonova-Catcher program is supported by the IdEx Université de Paris Cité ANR-18-IDEX-0001 and the MITI CNRS Sciences participative program.

Based on observations obtained with the Samuel Oschin Telescope 48 inch and the 60 inch Telescope at the Palomar Observatory as part of the Zwicky Transient Facility project. ZTF is supported by the National Science Foundation under grant No. AST-2034437 and a collaboration including Caltech, IPAC, the Weizmann Institute of Science, the Oskar Klein Center at Stockholm University, the University of Maryland, Deutsches Elektronen-Synchrotron and Humboldt University, the TANGO Consortium of Taiwan, the University of Wisconsin at Milwaukee, Trinity College Dublin, Lawrence Livermore National Laboratories, IN2P3, University of Warwick, Ruhr University Bochum and Northwestern University. Operations are conducted by COO, IPAC, and UW.

### ORCID iDs

Michael W. Coughlin  <https://orcid.org/0000-0002-8262-2924>  
 Joshua S. Bloom  <https://orcid.org/0000-0002-7777-216X>  
 Guy Nir  <https://orcid.org/0000-0002-7501-5579>  
 Sarah Antier  <https://orcid.org/0000-0002-7686-3334>  
 Theophile Jegou du Laz  <https://orcid.org/0009-0003-6181-4526>  
 Stéfan van der Walt  <https://orcid.org/0000-0001-9276-1891>  
 Arien Crellin-Quick  <https://orcid.org/0000-0002-7183-0410>  
 Thomas Culino  <https://orcid.org/0009-0008-3603-0013>  
 Dmitry A. Duev  <https://orcid.org/0000-0001-5060-8733>  
 Daniel A. Goldstein  <https://orcid.org/0000-0003-3461-8661>  
 Brian F. Healy  <https://orcid.org/0000-0002-7718-7884>  
 Viraj Karambelkar  <https://orcid.org/0000-0003-2758-159X>  
 Jada Lilleboe  <https://orcid.org/0009-0007-0764-7367>  
 Kyung Min Shin  <https://orcid.org/0000-0002-1486-3582>  
 Leo P. Singer  <https://orcid.org/0000-0001-9898-5597>  
 Tomás Ahumada  <https://orcid.org/0000-0002-2184-6430>  
 Shreya Anand  <https://orcid.org/0000-0003-3768-7515>  
 Eric C. Bellm  <https://orcid.org/0000-0001-8018-5348>  
 Richard Dekany  <https://orcid.org/0000-0002-5884-7867>  
 Matthew J. Graham  <https://orcid.org/0000-0002-3168-0139>

Mansi M. Kasliwal  <https://orcid.org/0000-0002-5619-4938>  
 Ivona Kostadinova  <https://orcid.org/0009-0003-6260-2891>  
 R. Weizmann Kiendrebeogo  <https://orcid.org/0000-0002-9108-5059>  
 Shrinivas R. Kulkarni  <https://orcid.org/0000-0001-5390-8563>  
 Sydney Jenkins  <https://orcid.org/0000-0001-9827-1463>  
 Natalie LeBaron  <https://orcid.org/0000-0002-2249-0595>  
 Ashish A. Mahabal  <https://orcid.org/0000-0003-2242-0244>  
 James D. Neill  <https://orcid.org/0000-0002-0466-1119>  
 B. Parazin  <https://orcid.org/0000-0002-3155-0385>  
 Julien Peloton  <https://orcid.org/0000-0002-8560-4449>  
 Daniel A. Perley  <https://orcid.org/0000-0001-8472-1996>  
 Reed Riddle  <https://orcid.org/0000-0002-0387-370X>  
 Ben Rusholme  <https://orcid.org/0000-0001-7648-4142>  
 Jakob van Santen  <https://orcid.org/0000-0002-2412-9728>  
 Jesper Sollerman  <https://orcid.org/0000-0003-1546-6615>  
 Robert Stein  <https://orcid.org/0000-0003-2434-0387>  
 D. Turpin  <https://orcid.org/0000-0003-1835-1522>  
 Avery Wold  <https://orcid.org/0000-0002-9998-6732>  
 Carla Amat  <https://orcid.org/0009-0008-6162-7882>  
 Adrien Bonnefon  <https://orcid.org/0009-0005-4020-4483>  
 Adrien Bonnefoy  <https://orcid.org/0009-0009-9174-0592>  
 Manon Flament  <https://orcid.org/0009-0005-0586-7281>  
 Frank Kerkow  <https://orcid.org/0009-0000-4242-7406>  
 Sulekha Kishore  <https://orcid.org/0009-0001-7793-3680>  
 Shloke Jani  <https://orcid.org/0009-0009-6530-4295>  
 Stephen K. Mahanty  <https://orcid.org/0009-0003-9879-7694>  
 Céline Liu  <https://orcid.org/0009-0001-9218-7196>  
 Laura Llinares  <https://orcid.org/0009-0008-8840-1576>  
 Jolyane Makarison  <https://orcid.org/0009-0008-9749-4604>  
 Alix Ollieric  <https://orcid.org/0009-0000-1435-9178>  
 Inès Perez  <https://orcid.org/0009-0008-8448-4773>  
 Lydie Pont  <https://orcid.org/0009-0000-2196-8481>  
 Vyom Sharma  <https://orcid.org/0009-0006-8436-3192>

### References

- Aartsen, M., Ackermann, M., Adams, J., et al. 2017, *JInst*, 12, P03012  
 Aasi, J., Abbott, B. P., Abbott, R., et al. 2015, *CQGra*, 32, 074001  
 Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017, *Natur*, 551, 85  
 Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017a, *PhRvL*, 119, 161101  
 Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017b, *ApJL*, 848, L12  
 Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017c, *ApJL*, 848, L13  
 Acernese, F., Agathos, M., Agatsuma, K., et al. 2015, *CQGra*, 32, 024001  
 Ackley, K., Amati, L., Barbieri, C., et al. 2020, *A&A*, 643, A113  
 Ahumada, T., Anand, S., Coughlin, M. W., et al. 2022, *ApJ*, 932, 40  
 Ahumada, T., Singer, L. P., Anand, S., et al. 2021, *NatAs*, 5, 917  
 Aivazyan, V., Almualla, M., Antier, S., et al. 2022, *MNRAS*, 515, 6007  
 Allan, A., Denny, R. B., & Swinbank, J. D. 2017, arXiv:1709.01264  
 Almualla, M., Coughlin, M. W., Anand, S., et al. 2020, *MNRAS*, 495, 4366  
 Anand, S., Coughlin, M. W., Kasliwal, M. M., et al. 2021, *NatAs*, 5, 46  
 Andreoni, I., Coughlin, M. W., Kool, E. C., et al. 2021, *ApJ*, 918, 63  
 Andreoni, I., Goldstein, D. A., Anand, S., et al. 2019, *ApJL*, 881, L16  
 Andreoni, I., Goldstein, D. A., Kasliwal, M. M., et al. 2020a, *ApJ*, 890, 131  
 Andreoni, I., Kool, E. C., Sagués Carracedo, A., et al. 2020b, *ApJ*, 904, 155  
 Antier, S., Agayeva, S., Aivazyan, V., et al. 2019, *MNRAS*, 492, 3904  
 Antier, S., Agayeva, S., Almualla, M., et al. 2020, *MNRAS*, 497, 5518  
 Arcavi, I., Hosseinzadeh, G., Howell, D. A., et al. 2017, *Natur*, 551, 64  
 Barbary, K., Barclay, T., Biswas, R., et al. 2016, SNCosmo: Python library for supernova cosmology, Astrophysics Source Code Library, ascl:1611.017  
 Barthelmy, S. D., Butterworth, P., Cline, T. L., et al. 1995, *Ap&SS*, 231, 235  
 Bellm, E. C., Kulkarni, S. R., Barlow, T., et al. 2019, *PASP*, 131, 068003  
 Biswas, B., Ishida, E. E. O., Peloton, J., et al. 2022, arXiv:2210.17433  
 Blagorodnova, N., Neill, J. D., Walters, R., et al. 2018, *PASP*, 130, 035003  
 Bloom, J. S., Starr, D. L., Butler, N. R., et al. 2008, *AN*, 329, 284  
 Bulla, M. 2019, *MNRAS*, 489, 5037  
 Bulla, M. 2023, *MNRAS*, 520, 2558

- Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv:1612.05560
- Chornock, et al. 2017, *ApJL*, **848**, L19
- Cook, D. O., Kasliwal, M. M., Van Sistine, A., et al. 2017, arXiv:1710.05016
- Coughlin, M. W. 2020, *NatAs*, **4**, 550
- Coughlin, M. W., Ahumada, T., Anand, S., et al. 2019a, *ApJL*, **885**, L19
- Coughlin, M. W., Ahumada, T., Cenko, S. B., et al. 2019b, *PASP*, **131**, 048001
- Coughlin, M. W., Antier, S., Corre, D., et al. 2019c, *MNRAS*, **489**, 5775
- Coughlin, M. W., Antier, S., Dietrich, T., et al. 2020b, *NatCo*, **11**, 4129
- Coughlin, M. W., Burdge, K., Duev, D. A., et al. 2021, *MNRAS*, **505**, 2954
- Coughlin, M. W., Dietrich, T., Antier, S., et al. 2019d, *MNRAS*, **492**, 863
- Coughlin, M. W., Dietrich, T., Antier, S., et al. 2020c, *MNRAS*, **497**, 1181
- Coughlin, M. W., Dietrich, T., Heinzel, J., et al. 2020a, *PhRvR*, **2**, 022006
- Coughlin, M. W., Tao, D., Chan, M. L., et al. 2018, *MNRAS*, **478**, 692
- Coulter, D. A., Foley, R. J., Kilpatrick, C. D., et al. 2017, *Sci*, **358**, 1556
- Coulter, D. A., Jones, D. O., McGill, P., et al. 2023, *PASP*, **135**, 064501
- Cowperthwaite, P. S., Berger, E., Villar, V. A., et al. 2017, *ApJL*, **848**, L17
- Cutri, R. M. 2012, *yCat*, **II/311**
- Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, 2MASS All Sky Catalog of Point Sources
- Dalya, G., Diaz, R., Bouchet, F., et al. 2021, *yCat*, **VII/291**
- De, K., Hankins, M. J., Kasliwal, M. M., et al. 2020, *PASP*, **132**, 025001
- Dekany, R., Smith, R. M., Riddle, R., et al. 2020, *PASP*, **132**, 038001
- Dietrich, T., Coughlin, M. W., Pang, P. T. H., et al. 2020, *Sci*, **370**, 1450
- Duev, D. A., Bolin, B. T., Graham, M. J., et al. 2021, *AJ*, **161**, 218
- Duev, D. A., Mahabal, A., Masci, F. J., et al. 2019, *MNRAS*, **489**, 3582
- Feindt, U., Nordin, J., Rigault, M., et al. 2019, *JCAP*, **2019**, 005
- Fernique, P., Boch, T., Donaldson, T., et al. 2014, arXiv:1505.02937
- Fernique, P., Nebot, A., Durand, D., et al. 2022, MOC: Multi-Order Coverage Map, IVOA Recommendation 27 July 2022, International Virtual Observatory Alliance
- Flewelling, H. 2018, AAS Meeting, **231**, 436.01
- Förster, F., Cabrera-Vives, G., Castillo-Navarrete, E., et al. 2021, *AJ*, **161**, 242
- Fremling, C., Hall, X. J., Coughlin, M. W., et al. 2021, *ApJL*, **917**, L2
- Fremling, C., Miller, A. A., Sharma, Y., et al. 2020, *ApJ*, **895**, 32
- Gehrels, N., Chincarini, G., Giommi, P., et al. 2004, *ApJ*, **611**, 1005
- Goldstein, A., Veres, P., Burns, E., et al. 2017, *ApJL*, **848**, L14
- Goldstein, D. A., Andreoni, I., Nugent, P. E., et al. 2019, *ApJL*, **881**, L7
- Gomez, S., Hosseinzadeh, G., Cowperthwaite, P. S., et al. 2019, *ApJL*, **884**, L55
- Gompertz, B. P., Cutter, R., Steeghs, D., et al. 2020, *MNRAS*, **497**, 726
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, **622**, 759
- Graham, M., Rixon, G., Dowler, P., et al. 2017, IVOA Support Interfaces Version 1.1, IVOA Recommendation 24 May 2017, International Virtual Observatory Alliance
- Graham, M. J., Kulkarni, S. R., Bellm, E. C., et al. 2019, *PASP*, **131**, 078001
- Guillochon, J., Nicholl, M., Villar, V. A., et al. 2018, *ApJS*, **236**, 6
- Ho, A. Y. Q., Perley, D. A., Beniamini, P., et al. 2020, *ApJ*, **905**, 98
- Hotokezaka, K., Nakar, E., Gottlieb, O., et al. 2019, *NatAs*, **3**, 940
- Ivezic, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, **873**, 111
- Kasliwal, M. M., Anand, S., Ahumada, T., et al. 2020, *ApJ*, **905**, 145
- Kasliwal, M. M., Kasen, D., Lau, R. M., et al. 2019, *MNRAS*, **510**, L7
- Kumar, R., Carroll, C., Hartikainen, A., & Martin, O. 2019, *JOSS*, **4**, 1143
- Li, W., Filippenko, A. V., Chornock, R., & Jha, S. 2003, *PASP*, **115**, 844
- Licorish, S. A., Galster, M., Kapitsaki, G. M., & Tahir, A. 2022, *J. Syst. Softw.*, **186**, 111156
- Lourie, N. P., Baker, J. W., Burruss, R. S., et al. 2020, *Proc. SPIE*, **11447**, 114479K
- Lundquist, M. J., Paterson, K., Fong, W., et al. 2019, *ApJL*, **881**, L26
- Mahabal, A., Rebbapragada, U., Walters, R., et al. 2019, *PASP*, **131**, 038002
- Masci, F. J., Laher, R. R., Rusholme, B., et al. 2019, *PASP*, **131**, 018003
- Matheson, T., Stubens, C., Wolf, N., et al. 2021, *AJ*, **161**, 107
- Meegan, C., Lichti, G., Bhat, P. N., et al. 2009, *ApJ*, **702**, 791
- Metzger, B. D. 2017, *LRR*, **20**, 3
- Möller, A., Peloton, J., Ishida, E. E. O., et al. 2020, *MNRAS*, **501**, 3272
- Naul, B., van der Walt, S., Crellin-Quick, A., Bloom, J. S., & Pérez, F. 2016, arXiv:1609.04504
- Nordin, J., Brinnel, V., van Santen, J., et al. 2019, *A&A*, **631**, A147
- Pang, P. T. H., Dietrich, T., Coughlin, M. W., et al. 2022, arXiv:2205.08513
- Parazin, B., Coughlin, M. W., Singer, L. P., Gupta, V., & Anand, S. 2022, *ApJ*, **935**, 87
- Patterson, M. T., Bellm, E. C., Rusholme, B., et al. 2018, *PASP*, **131**, 018001
- Perley, D. A., Fremling, C., Sollerman, J., et al. 2020, *ApJ*, **904**, 35
- Petrov, P., Singer, L. P., Coughlin, M. W., et al. 2022, *ApJ*, **924**, 54
- Pian, E., D'Avanzo, P., Benetti, S., et al. 2017, *Natur*, **551**, 67
- Reichherzer, P., Schüssler, F., Lefranc, V., et al. 2021, *ApJS*, **256**, 5
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, *Proc. SPIE*, **9143**, 914320
- Rigault, M., Neill, J. D., Blagorodnova, N., et al. 2019, *A&A*, **627**, A115
- Rosswog, S., Feindt, U., Korobkin, O., et al. 2017, *CQGra*, **34**, 104001
- Rosswog, S., Sollerman, J., Feindt, U., et al. 2018, *A&A*, **615**, A132
- Ryan, G., Eerten, H. v., Piro, L., & Troja, E. 2020, *ApJ*, **896**, 166
- Savchenko, V., Ferrigno, C., Kuulkers, E., et al. 2017, *ApJL*, **848**, L15
- Seaman, R., Williams, R., Allan, A., et al. 2011, Sky Event Reporting Metadata Version 2.0, IVOA Recommendation 11 July 2011, International Virtual Observatory Alliance
- Singer, L., & Racusin, J. 2023, AAS Meeting, **241**, 108.02
- Singer, L. P., Parazin, B., Coughlin, M. W., et al. 2022, *AJ*, **163**, 209
- Smartt, S. J., Chen, T.-W., Jerkstrad, A., et al. 2017, *Natur*, **551**, 75
- Smith, K. W., Smartt, S. J., Young, D. R., et al. 2020, *PASP*, **132**, 085002
- Smith, K. W., Williams, R. D., Young, D. R., et al. 2019, *RNAAS*, **3**, 26
- Steele, I. A., Smith, R. J., Rees, P. C., et al. 2004, *Proc. SPIE*, **5489**, 679
- Stonebraker, M., & Rowe, L. A. 1986, in Proc. of the 1986 ACM SIGMOD Int. Conf. on Management of Data, SIGMOD '86, ed. C. Zaniolo (New York: Association for Computing Machinery), 340
- Street, R. A., Bowman, M., Saunders, E. S., & Boroson, T. 2018, *Proc. SPIE*, **10707**, 1070711
- Tachibana, Y., & Miller, A. A. 2018, *PASP*, **130**, 128001
- Tonry, J. L., Denneau, L., Heinze, A. N., et al. 2018, *PASP*, **130**, 064505
- Valenti, S., Sand, D. J., Yang, S., et al. 2017, *ApJL*, **848**, L24
- van der Walt, S. J., Crellin-Quick, A., & Bloom, J. S. 2019, *JOSS*, **4**, 1247
- van Roestel, J., Duev, D. A., Mahabal, A. A., et al. 2021, *AJ*, **161**, 267
- Watson, D., Hansen, C. J., Selsing, J., et al. 2019, *Natur*, **574**, 497
- Wyatt, S. D., Tohuvavohu, A., Arcavi, I., et al. 2020, *ApJ*, **894**, 127
- Yao, Y., Gezari, S., Velzen, S. V., et al. 2021, *TNSAN*, **2021-258**
- Zechmeister, M., & Kürster, M. 2009, *A&A*, **496**, 577