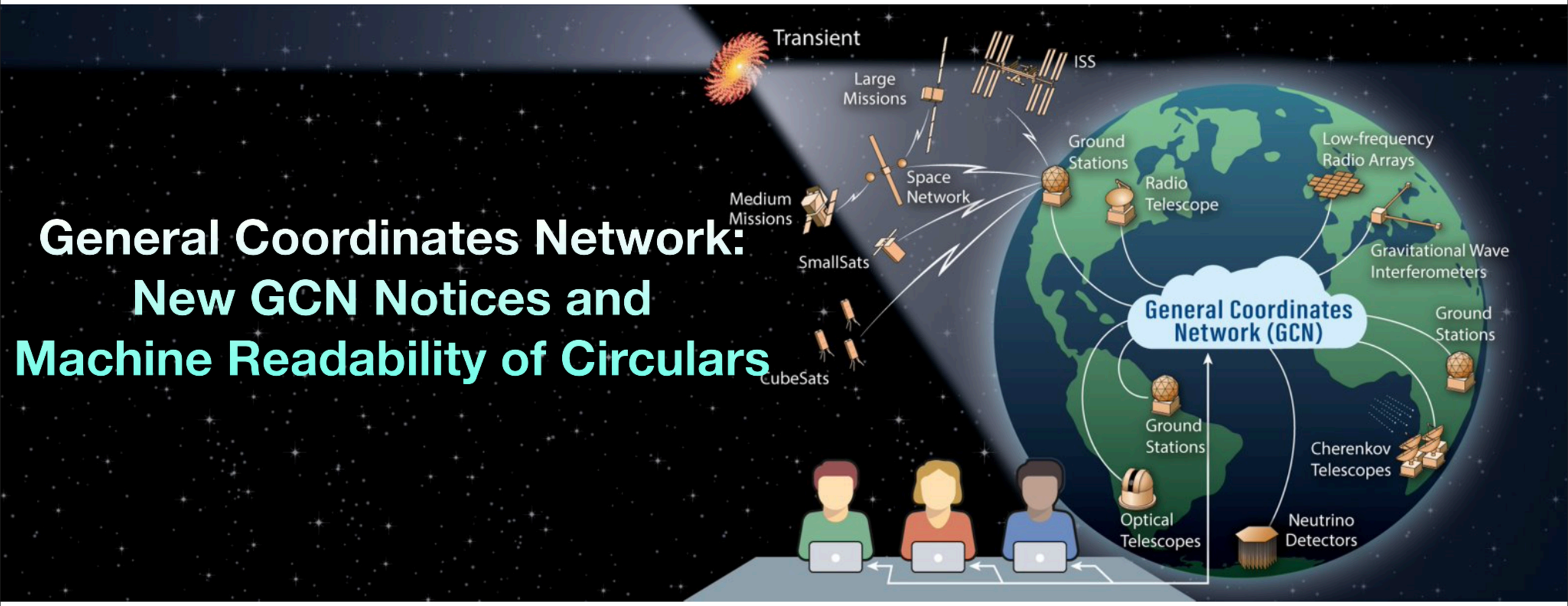
Astro-COLIBRI Multimessenger Astrophysics workshop
September 17th, 2024

National Aeronautics and Space Administration

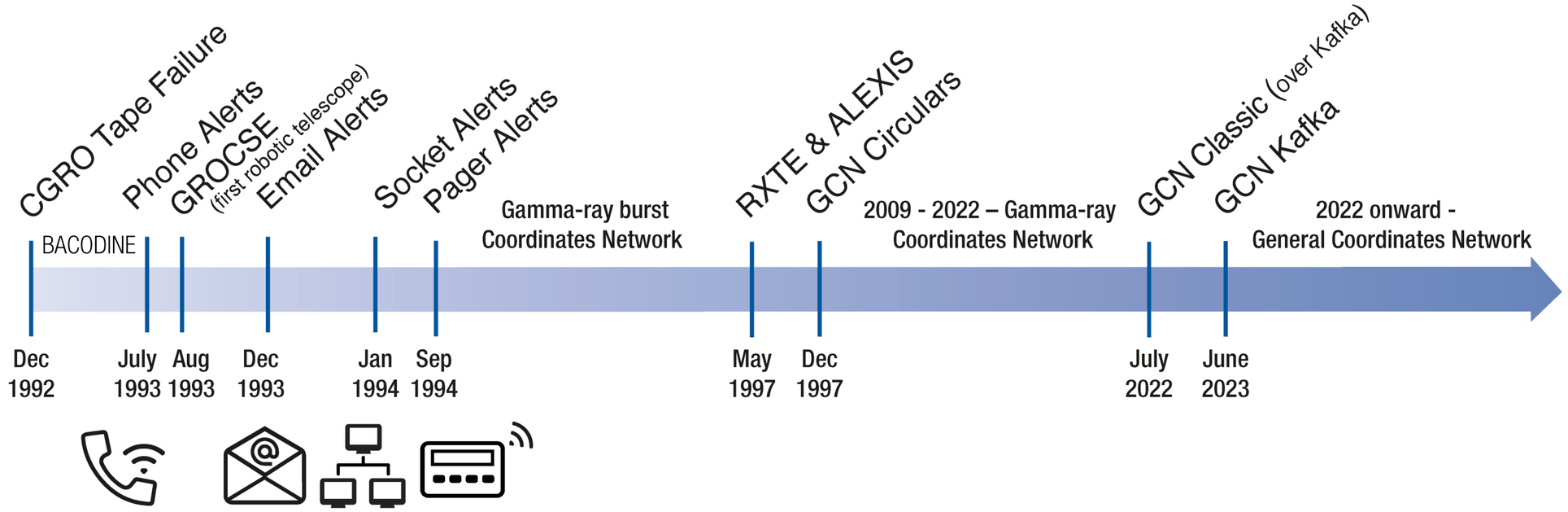
General Coordinates Network: New GCN Notices and Machine Readability of Circulars



Vidushi Sharma & GCN Team
NASA GSFC/UMBC

https://gcn.nasa.gov

Early History of General Coordinates Network (GCN)



- Compton Gamma-Ray Observatory's (CGRO) onboard recorder failed in 1992
- Need to downlink events as they occurred created an opportunity for realtime follow-up
- BAtse COordinates DIstribution NEtwork (BACODINE) was built to receive and distribute those alerts worldwide
- BACODINE provided new alert formats (phone, email, socket, and pager)
- New instruments and transient types led to the Gamma-ray Coordinates Network

There are two kinds of GCN data products:

GCN NOTICES

```
TITLE:            GCN/FERMI NOTICE
NOTICE_DATE:      Wed 26 Aug 20 22:10:07 UT
NOTICE_TYPE:      Fermi-GBM Flight Position
RECORD_NUM:       45
TRIGGER_NUM:      620172587
GRB_RA:           296.300d {+19h 45m 12s} (J2000),
                  296.250d {+19h 45m 00s} (current),
                  296.416d {+19h 45m 40s} (1950)
GRB_DEC:          +71.817d {+71d 49' 00"} (J2000),
                  +71.868d {+71d 52' 03"} (current),
                  +71.693d {+71d 41' 35"} (1950)
GRB_ERROR:        5.50 [deg radius, statistical plus systematic]
GRB_INTEN:        1078 [cnts/sec]
DATA_SIGNIF:      22.80 [sigma]
INTEG_TIME:       1.024 [sec]
GRB_DATE:         19087 TJD;   239 DOY;   20/08/26
GRB_TIME:         79782.72 SOD {22:09:42.72} UT
GRB_PHI:           20.00 [deg]
GRB_THETA:        150.00 [deg]
DATA_TIME_SCALE: 1.0240 [sec]
HARD_RATIO:       0.54
LOC_ALGORITHM:    3 (version number of)
MOST_LIKELY:       93%  GRB
2nd_MOST_LIKELY:    4%  Generic Transient
DETECTORS:        0,0,0, 0,1,1, 0,0,0, 0,0,0, 0,0,
SUN_POSTN:        156.00d {+10h 24m 01s}  +10.00d {+09d 59' 51"}
SUN_DIST:          94.05 [deg]   Sun_angle= -9.3 [hr] (East of Sun)
MOON_POSTN:       258.31d {+17h 13m 14s}  -22.27d {-22d 15' 56"}
MOON_DIST:         97.64 [deg]
MOON_ILLUM:       63 [%]
GAL_COORDS:       103.87, 21.63 [deg] galactic lon,lat of the burst (or transient)
ECL_COORDS:        41.25, 79.40 [deg] ecliptic lon,lat of the burst (or transient)
LC_URL:           http://heasarc.gsfc.nasa.gov/FTP/fermi/data/gbm/triggers/2020/bn200826923/
COMMENTS:         Fermi-GBM Flight-calculated Coordinates.
COMMENTS:         This trigger occurred at longitude,latitude = 209.65,1.28 [deg].
COMMENTS:         The LC_URL file will not be created until ~15 min after the trigger.
```

- By and for machines
- Fixed, predefined format
- Schema specific to each notice type

GCN CIRCULARS

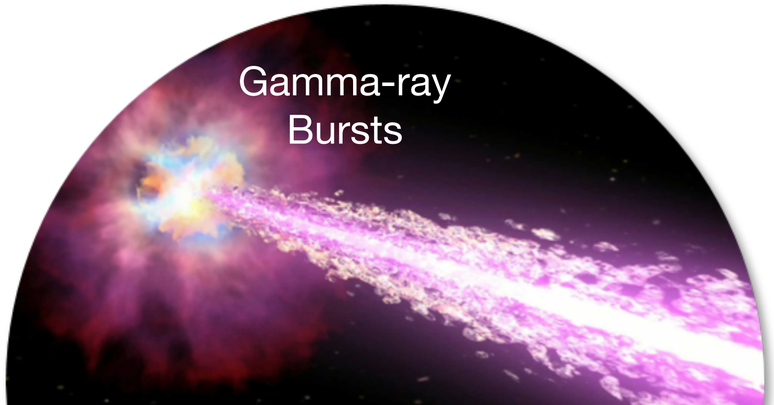
```
TITLE:   GCN CIRCULAR
NUMBER:  28298
SUBJECT: GRB 200826B: Fermi GBM detection
DATE:    20/08/27 21:10:30 GMT
FROM:    Christian Malacaria at NASA-MSFC/USRA  <cmalacaria@usra.edu>

C. Malacaria (NASA-MSFC/USRA) and C.Meegan (UAH)
report on behalf of the Fermi GBM Team:

"At 22:09:42.72 UT on 26 August 2020, the Fermi Gamma-Ray Burst Monitor (GBM)
triggered and located GRB 200826B (trigger 620172587 / 200826923).

The on-ground calculated location, using the GBM trigger
data, was reported in GCN 28292.

The GBM light curve shows an exceptionally bright long GRB
with a duration (T90) of about 7.4 s (50-300 keV).
The time-averaged spectrum from T0-0.003 s to T0+ 12.544 s is
best fit by a Band function with Epeak = 410.3 +/- 5.6  keV,
alpha = -0.64 +/- 0.01, and beta = -2.52 +/- 0.04
The event fluence (10-1000 keV) in this time interval is
(1.414 +/- 0.006)E-04 erg/cm^2.
The 1.024-sec peak photon flux measured starting from T0+5.1 s in
the 10-1000 keV band is 110.1 +/- 0.7 ph/s/cm^2.

The spectral analysis results presented above are preliminary;
final results will be published in the GBM GRB Catalog:
https://heasarc.gsfc.nasa.gov/W3Browse/fermi/fermigbrst.html

For Fermi GBM data and info, please visit the official Fermi GBM Support Page:
https://fermi.gsfc.nasa.gov/ssc/data/access/gbm/"
```

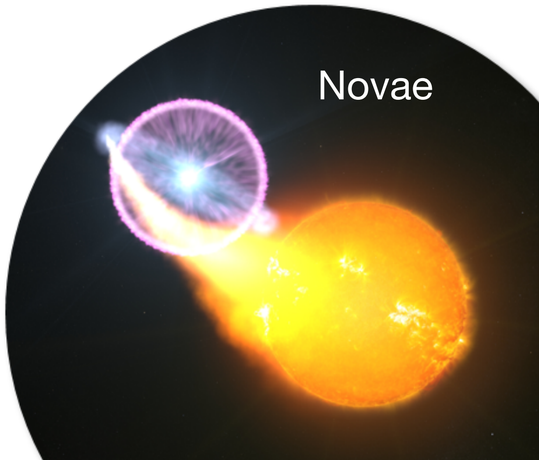
- By and for humans (some automated)
- Freeform text (with established style)
- Citable (but not peer-reviewed)

Changing Scientific Landscape: Fast Astronomical Transients

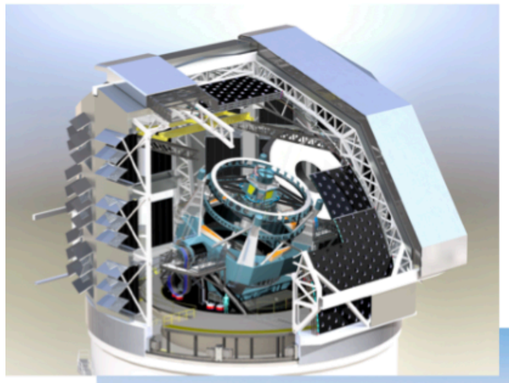
GCN is constantly evolving to serve new transients, messengers, and observatories:

- Gravitational wave events (GW150914, GW170817)
- High-energy neutrinos (IC170922A)
- Tidal disruption events (Swift J1644+57)
- Magnetar giant flares (200415A)





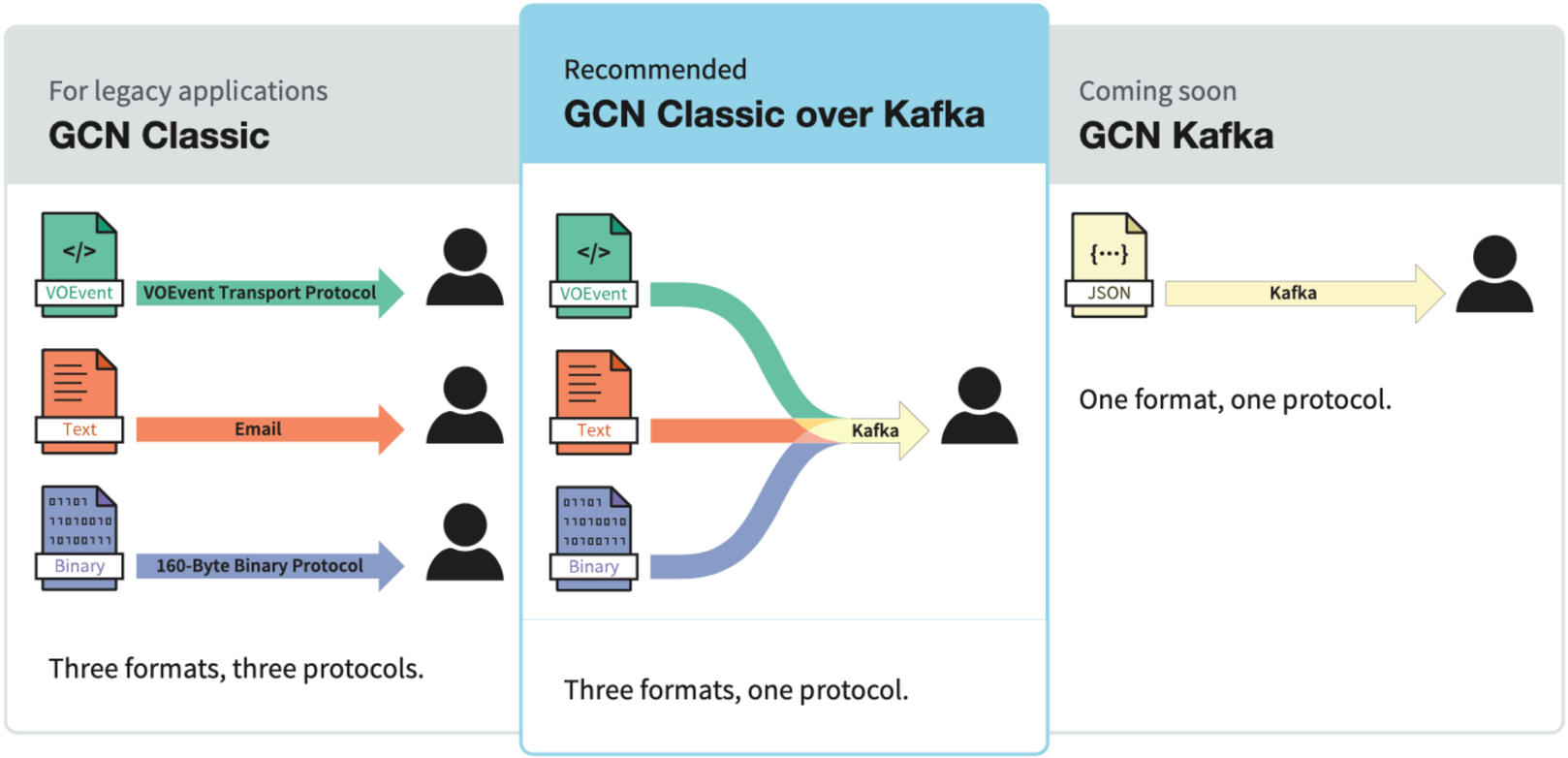
Changing Technological Landscape



Rubin Observatory/NSF/AURA

- Internet standards have led to new, better ways to serialize astronomy data (VOEvent, JSON, Avro, etc.)
- Encryption is necessary on the modern Internet (e.g. https)
- Industry has developed general time-series databases and streaming frameworks
- The Vera C. Rubin Observatory will use Apache Kafka to distribute transient alerts as its primary data product
- Many other experiments are following suit: Zwicky Transient Facility, LIGO/Virgo/KAGRA

Introducing the New GCN: Built on Kafka



- GCN Classic provides three formats over *three custom protocols*
- GCN Classic over Kafka provides all three formats over *one standard protocol*: Apache Kafka

New GCN is built on standard protocol for streaming alerts: ***Apache Kafka***

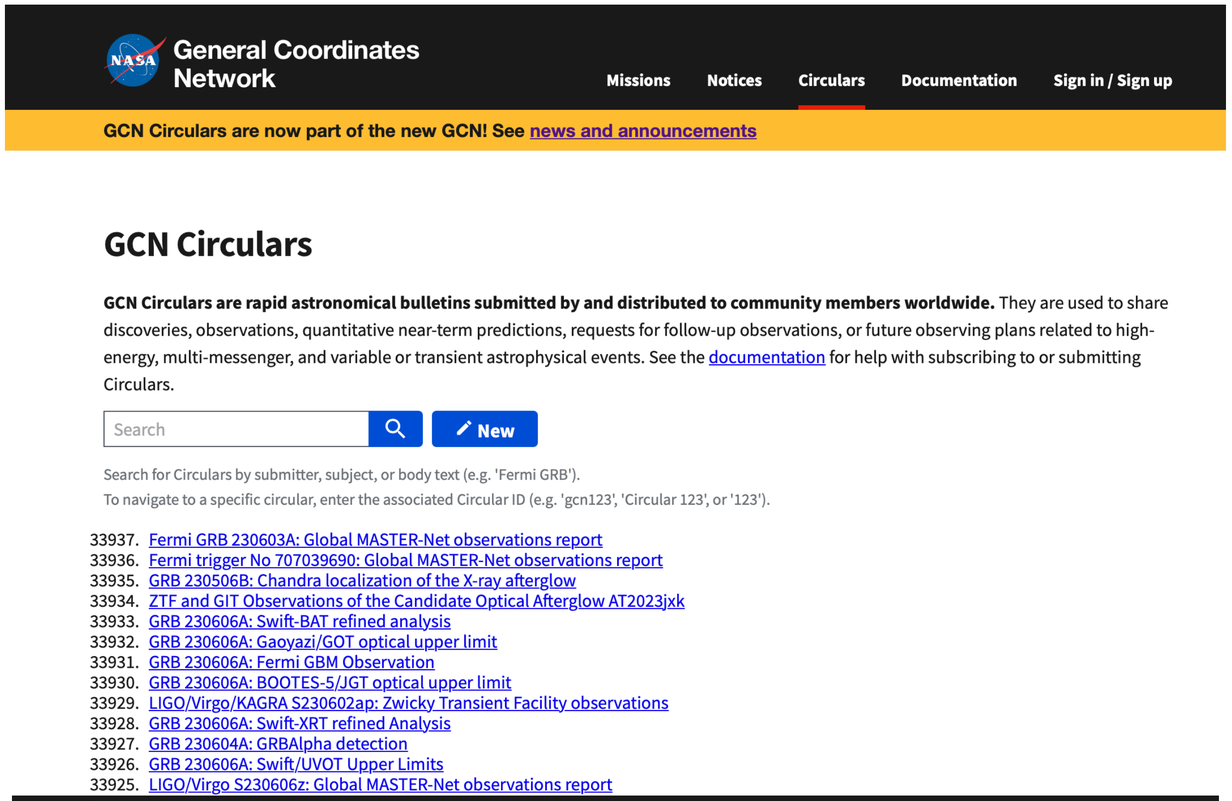
Why to Switch to New GCN?

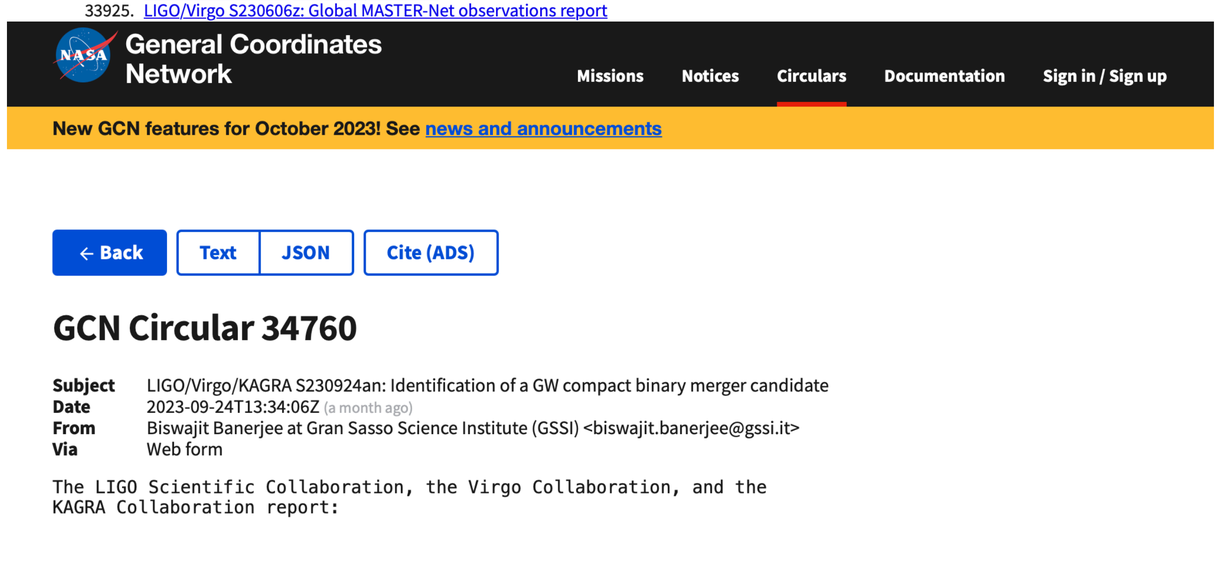
	GCN Classic	GCN Classic over Kafka
Self-service	**NO.** Users need to contact administrator in order to make account and subscription changes	**YES.** Manage your own account and subscription settings through the web site
Open standards	**NO.** Notices are sent using three custom protocols	**YES.** Notices are sent using one standard protocol, Apache Kafka
Open source	**NO.** Custom software needed to receive notices	**YES.** Receive notices using open-source software
Highly available	**NO.** Notices are broadcast by a single server	**YES.** Notices are broadcast by a cluster of highly-available Kafka brokers in the cloud
Secure	**NO.** Notices are sent as plaintext	**YES.** Notices are protected with SSL/TLS

New GCN Website

at https://gcn.nasa.gov

- Updated look and feel
- More accessible, based on US Web Design System
- Single sign on with:
  - Email and password
  - Google
  - Facebook
  - LaunchPad (for NASA employees and affiliates)





New and improved:

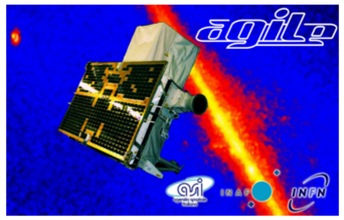
✨GCN Circulars✨

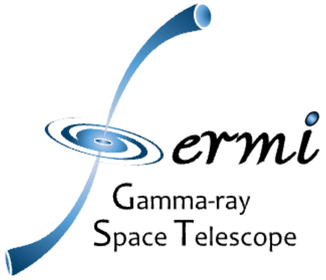
at https://gcn.nasa.gov/circulars

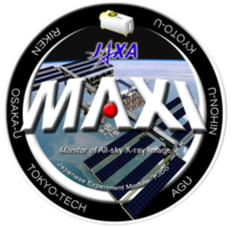
- Browse and search our new archive.
- Manage your own email subscriptions.
- Enroll yourself and your colleagues to submit Circulars with arXiv-style peer endorsements.
- Submit Circulars with our new Web form, or continue to submit by email.
- Real-time integration with SAO/NASA Astrophysics Data Service (ADS)

GCN Classic Notices

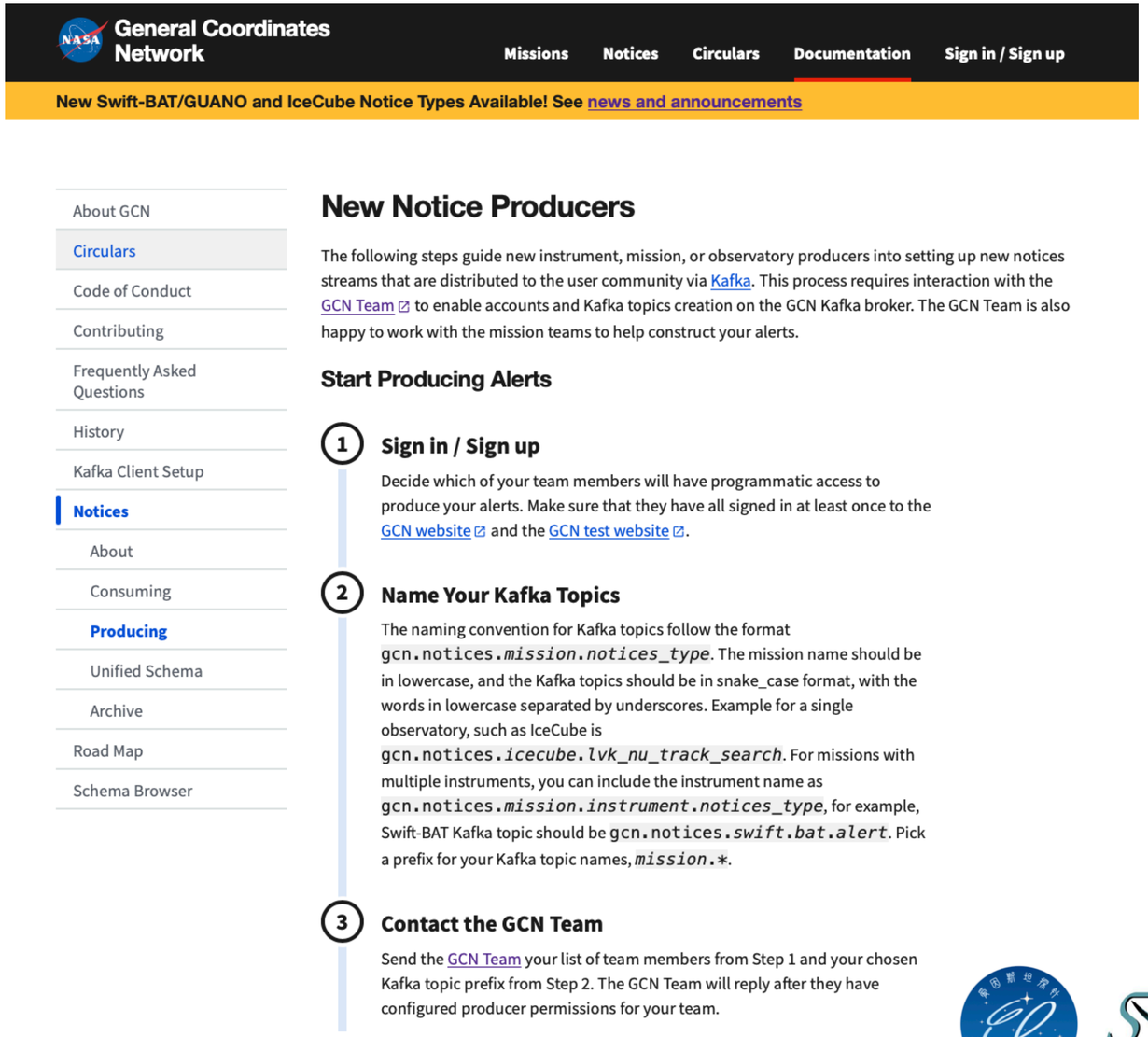
GCN Notice Producer Missions/Observatories/Experiments







Create New GCN Notices



- New Notices topics streamed by only GCN Kafka
- **For step by step instructions: Notices > Producing**
- New Notices produced by only Unified Schema: IceCube, Swift-BAT Guano are our first new Notices producers
- **Notices format: JSON**
- Create your Kafka Topic
- **Draft your Schema**

ICECUBE
GW Follow-up

New GCN Notices

New Notice Types: Unified schema and alert format

- JSON Core Schema with common core fields and consistent units
- Instrument/mission/observatory specific fields where needed
- https://github.com/nasa-gcn/gcn-schema
- New Producers: IceCube, Swift/BAT-GUANO, Einstein Probe
- BurstCube, SVOM, SGR and gamma-ray transients New Notices by IPN (soon)

JSON example of IceCube GW Follow-up Schema:

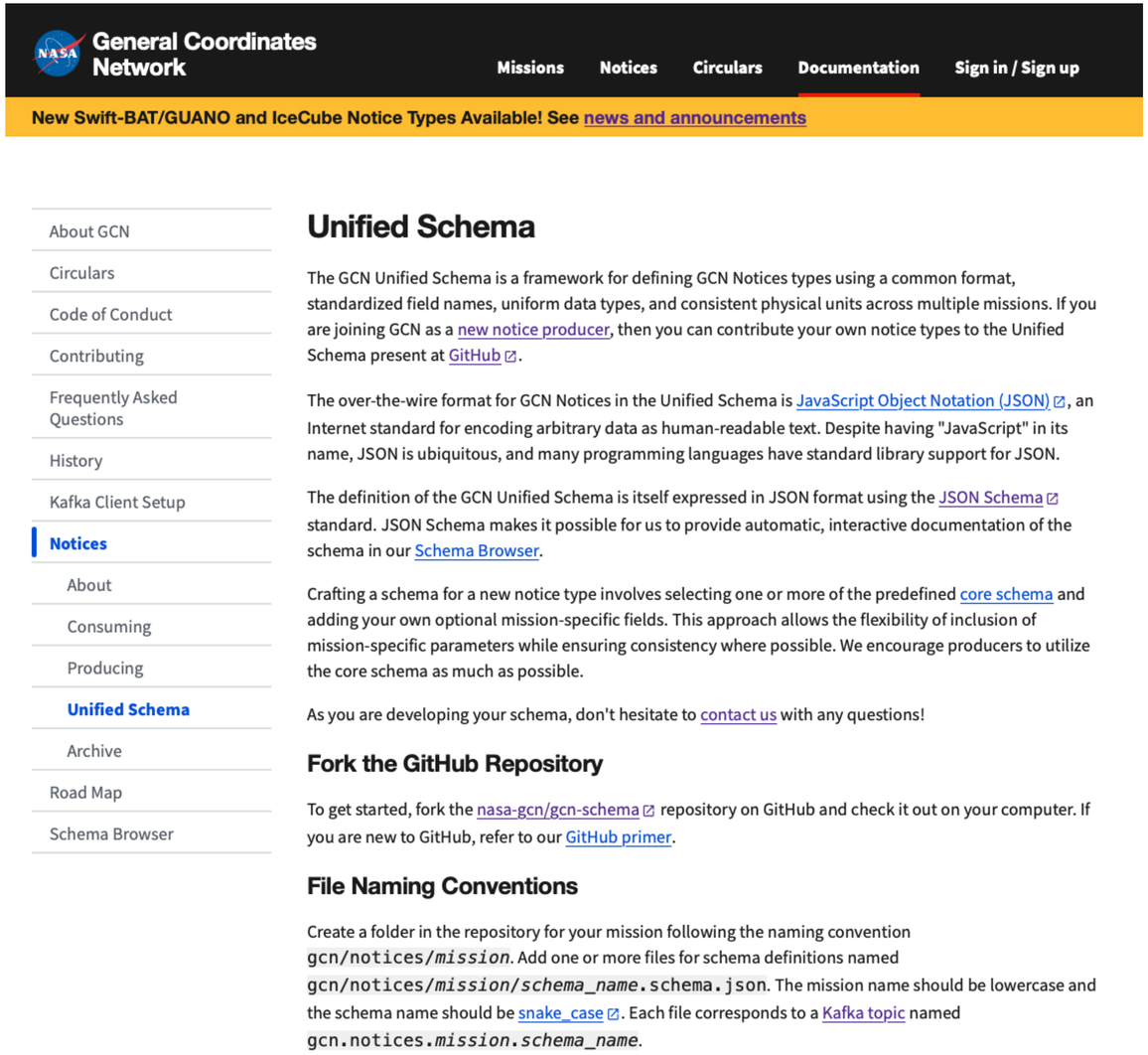
Schema


```
{
  "$id": "https://gcn.nasa.gov/schema/gcn/notices/icecube/LvkNuTrackSearch.schema.json",
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "type": "object",
  "title": "IceCube LVK Alert Nu Track Search",
  "description": "IceCube LVK Coincident Neutrino Track Search",
  "allOf": [
    {
      "$ref": "/schema/gcn/notices/core/Alert.schema.json",
      "description": "Alert information (alert_datetime and alert_type) from LVK alert notices"
    },
    {
      "$ref": "/schema/gcn/notices/core/FollowUp.schema.json",
      "description": "ID (ref_ID) of LVK alert notice"
    },
    {
      "$ref": "/schema/gcn/notices/core/AdditionalInfo.schema.json"
    },
    {
```

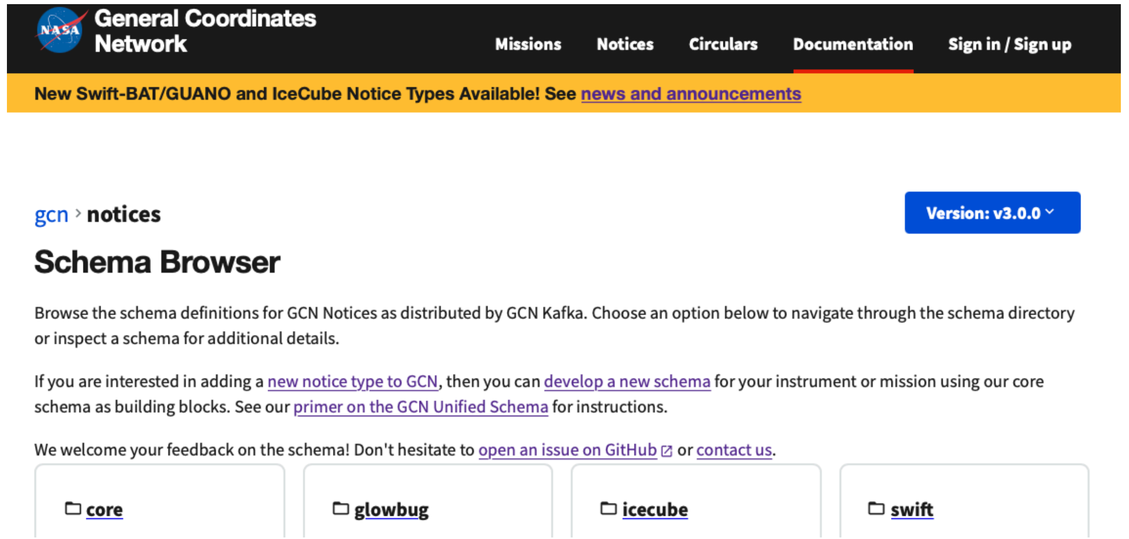

Notice


```
{
  "$schema": "https://gcn.nasa.gov/schema/gcn/notices/icecube/LvkNu
  "type": "IceCube LVK Alert Nu Track Search",
  "reference": { "gcn.notices.LVK.alert": 4642 },
  "ref_ID": "MS230427r",
  "alert_datetime": "2023-04-16T05:32:29.55Z",
  "trigger_time": "2023-04-16T05:22:26.0Z",
  "observation_start": "2023-04-16T05:14:06.0Z",
  "observation_stop": "2023-04-16T05:30:46.0Z",
  "observation_livetime": 1000,
  "pval_generic": 0.5,
  "pval_bayesian": 0.45,
  "n_events_coincident": 2,
  "coincident_events": [
    {
      "event_dt": -123.4,
      "ra": 345.82,
```

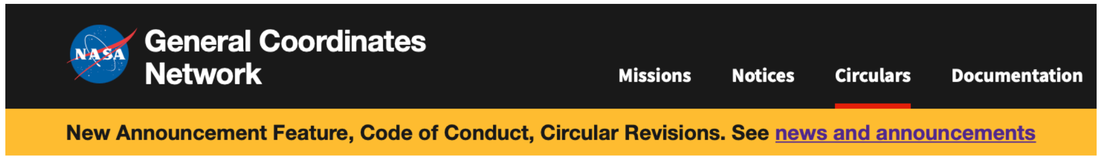
Create Mission Schema



- Fork & Set-up GitHub Repository:
  - GCN Schema v3.0.0
- Design Your Schema
  - See Sample code/Existing Examples
  - Use Core Schema & Specific fields
  - Validate & Submit Schema for Feedback
- Explore Schema-Browser for Schema definitions and examples



GCN Circulars: Large Language Model application



← Back | Text | JSON | Cite

GCN Circular 21520

Subject GRB 170817A: Fermi GBM detection
Date 2017-08-17T20:00:07Z (7 years ago)
From Andreas von Kienlin at MPE <azk@mpe.mpg.de>

A. von Kienlin (MPE), C. Meegan (UAH) and A. Goldstein (USRA)
report on behalf of the Fermi GBM Team:

"At 12:41:06.47 UT on 17 August 2017, the Fermi Gamma-Ray Burst Monitor
triggered and located GRB 170817A (trigger 524666471 / 170817529).

The on-ground calculated location, using the GBM trigger
data, is RA = 176.8, DEC = -39.8 (J2000 degrees,
equivalent to 12 h 47 m, -39 d 48'), with an uncertainty
of 11.6 degrees (radius, 1-sigma containment,
statistical only; there is additionally a systematic
error which we have characterized as a core-plus-tail model, with 90% of
GRBs having a 3.7 deg error and a small tail suffering a larger than 10 deg
systematic error. [Connaughton et al. 2015, ApJS, 216, 32]).

The angle from the Fermi LAT boresight at the GBM trigger time is 91 degrees.

The GRB light curve shows a weak short pulse
with a duration (T90) of about 2 s (50-300 keV).
The time-averaged spectrum from T0-0.512 s to 2.048 s is
well fit by a power law function with an exponential
high-energy cutoff. The power law index is -0.89 +/- 0.5 and
the cutoff energy, parameterized as Epeak, is 82 +/- 21 keV

The event fluence (10-1000 keV) in this time interval is
(2.3 +/- 0.4)E-07 erg/cm^2. The 1.024-sec peak photon flux measured
starting from T0-0.32 s in the 8-1000 keV band
is 1.9 +/- 0.2 ph/s/cm^2.

The spectral analysis results presented above are preliminary;
final results will be published in the GBM GRB Catalog."

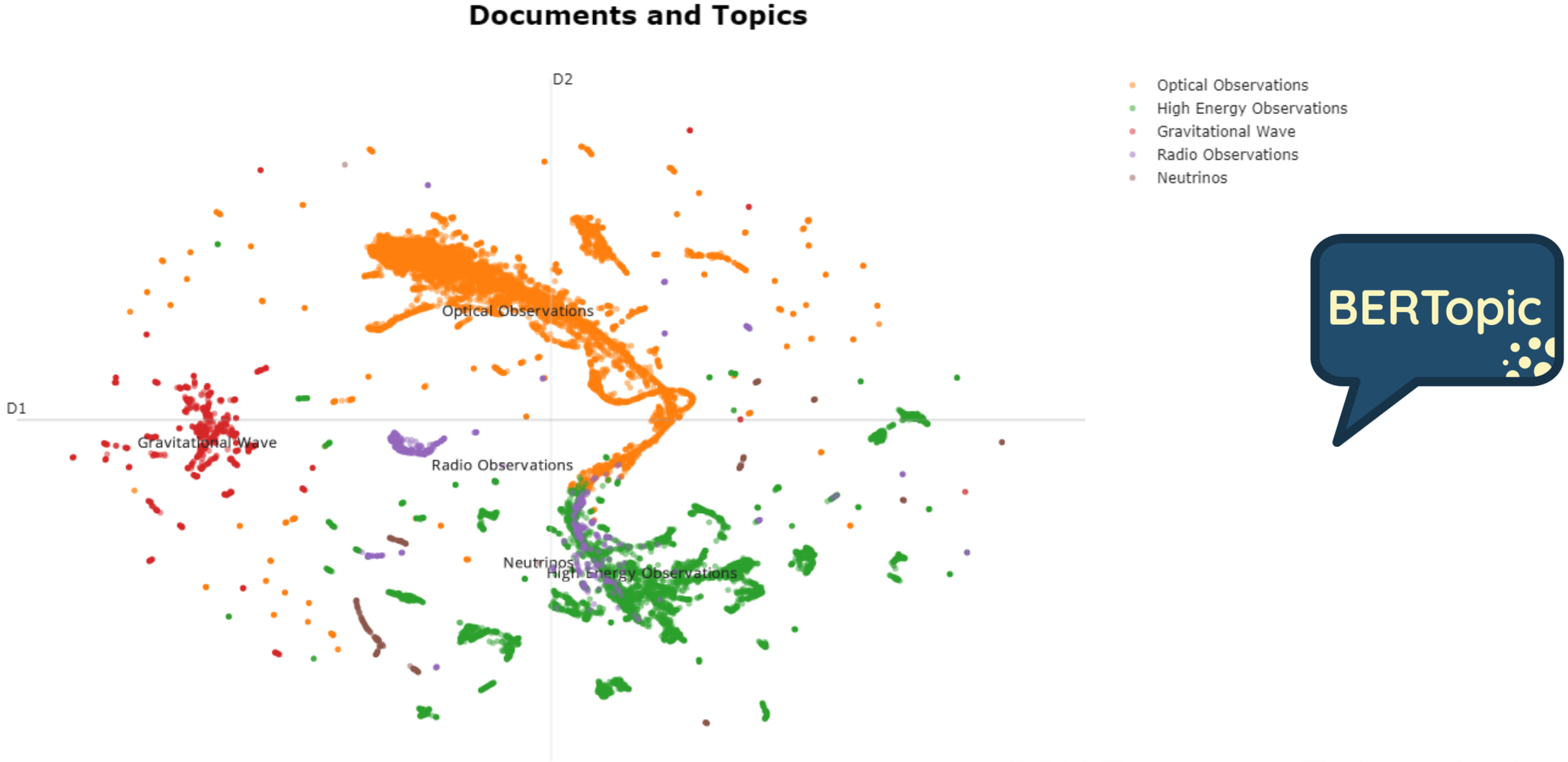
- ✦ Human-written, Flexible and Unstructured alerts
- ✦ Difficult to parse with conventional methods

LLM trained for human-written text:

- Topic modeling, such as observation-based cluster
- Information Extraction, such as Redshift

GCN Circulars: Topic Modeling for observation-type clustering

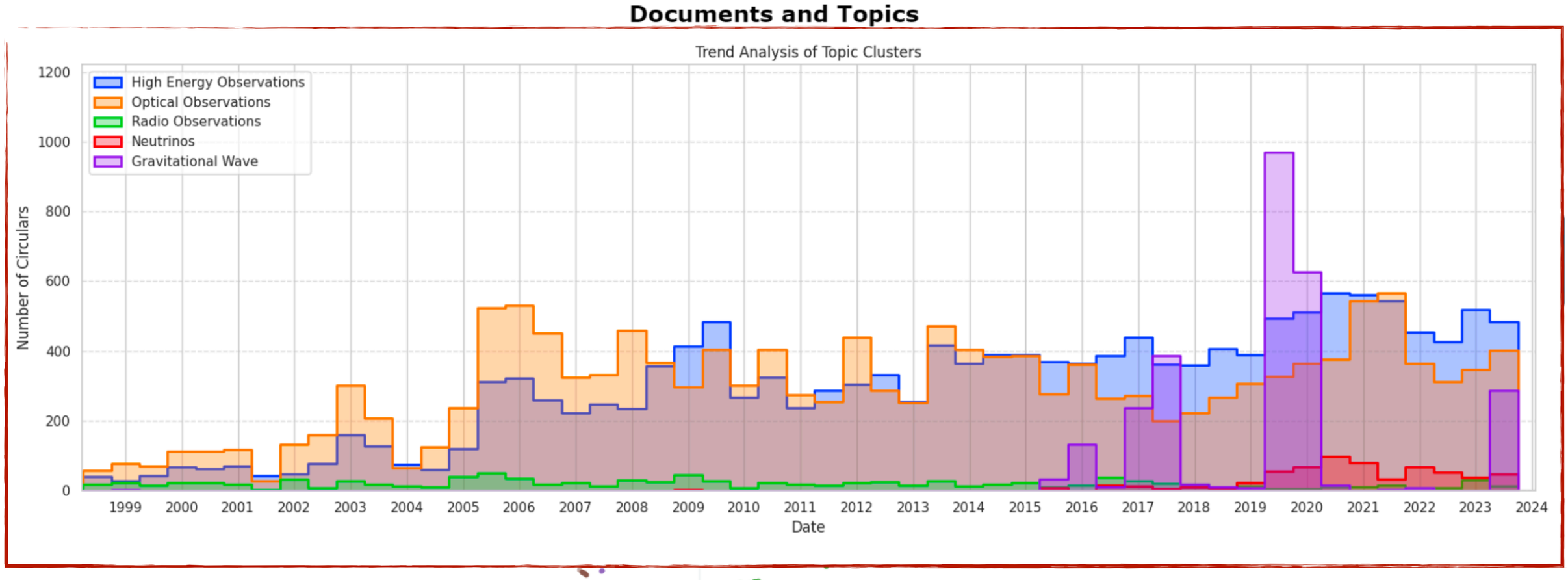
- BERT (Bidirectional Encoder Representations from Transformers) - Google AI
- BERT architecture based model “all-MiniLM-L6-v2” - fine-tuned for observational based clustering
- 5 Different type of observational cluster are extracted with BERTopic library



GCN Team et al (To be submitted)

GCN Circulars: Topic Modeling for observation-type clustering

- BERT (Bidirectional Encoder Representations from Transformers) - Google AI
- BERT architecture based model “all-MiniLM-L6-v2” - fine-tuned for observational based clustering
- 5 Different type of observational cluster are extracted with BERTopic library



GCN Circulars: Information Extraction

Mistral-7B-Instruct-v0.2

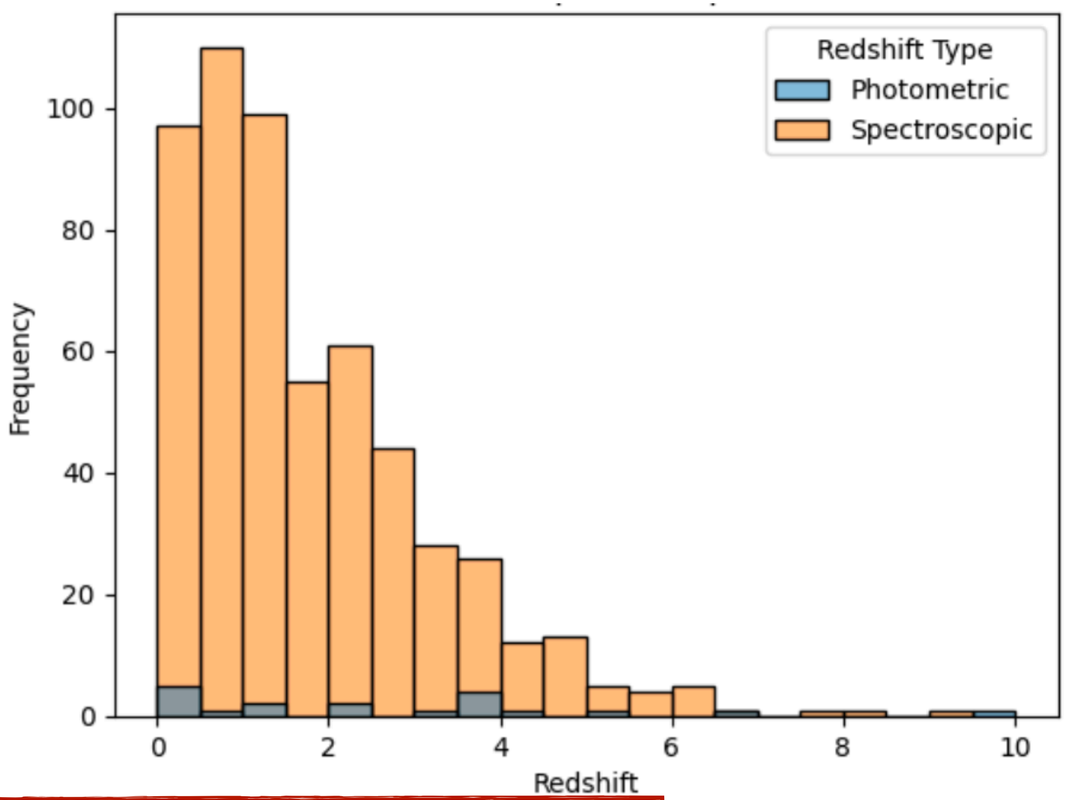
-Fine-tuned generative text-model

-Trained on publicly available conversational data-search

-Prompt Tuning output parsing and RAG done with LangChain library

-Info: Redshift values, event names, instrument, redshift-type

-Accuracy measured on Swift GRB table is 98%



GCN	Actual z	Predicted z	Actual GRB	Predicted GRB	Actual Telescope	Predicted Telescope
9457	2.625	z = 2.63	090529	GRB090529	VLT	ESO Very Large Telescope
9518	0.54	z = 0.54	090618	GRB090618	Lick	Lick Observatory
9542	0.54	0.54	090618	GRB090618	SAO RAS	SAO RAS
9673	3.00	3.00	090715B	GRB090715B	WHT	William Herschel Telescope
9712	2.71	z = 2.71	090726	GRB090726	SAO RAS	SAO RAS 6-m telescope
9761	2.737	2.737 +/- 0.002	090809	GRB090809	VLT	Kueyen telescope of ESO's VLT
9771	2.452	2.452	090812	GRB090812	VLT	VLT (Paranal observatory)

GCN Team et al (To be submitted)

Thanks for listening!

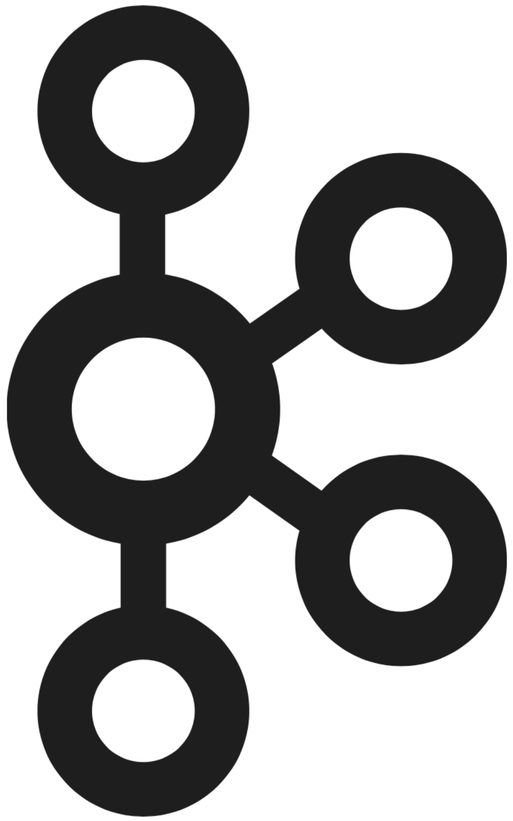
Web site: https://gcn.nasa.gov
This presentation: https://nasa-gcn.github.io/gcn-presentation/

Questions or comments? Contact GCN directly

Have you found a bug in GCN? Open an issue

Want to contribute code to GCN? Get involved on GitHub

What is Kafka?



Apache Kafka is an open-source distributed event streaming platform used by thousands of companies for high-performance data pipelines, streaming analytics, data integration, and mission-critical applications.

— from https://kafka.apache.org/

Kafka is widely used at NASA

- Existing Kafka applications at NASA include:
  - GCN (Goddard Space Flight Center)
  - Complex Event Processor - Deep Space Network (Jet Propulsion Laboratory)
  - Enterprise Business Information Services (Jet Propulsion Laboratory)
  - Federated Airspace Management Framework (Ames Research Center)
- All Federal agencies are using self-managed Kafka brokers, either Apache Kafka or Confluent Platform
- GCN is sponsoring FedRAMP authorization for Confluent Cloud to make it easy for NASA and other federal agencies to deploy Kafka software-as-a-service