

# GPT4o data extraction from ATels

Lizabeth Joy Tan, Musfar Muhamed, Andrei Variu

# Goals

- Create AstroCOLIBRI alerts from ATels
- Create a DB for sources:
  - Provide all ATels of a given source
- Can GPT4o suggest actions based on ATel/GCN:
  - E.g. take a light-curve from ASASSN
  - Checked a bit:
    - Input prompt: a list of things Astro-COLIBRI can do and an ATel indicating a transient was observed
    - When ask what's the best thing to do given the alert, GPT output the first thing from the list or outputs them in order of what is listed in the prompt
  - Does not seem to work

# Data extraction

- <https://gitlab.renkulab.io/andrei.variu/astrocolibri-gpt-data-extraction> (and forks)
- Tried “Agents”; does not seem to work
- Prompt:

Extract information from the text and provide the output in JSON format as:

```
{"Celestial Object":  
  [{"Name": , "RA":, "Dec":, "redshift":, "Type":, "Time":, "Date": }],  
  "References": ,  
  "Event Type":,  
  "Time":,  
  "Date":,  
  "Telescope": ,  
  "Instrument":,  
  "Observatory":}
```

for all astronomical object names and put Null when there is not information.

<<ATEL TEXT>>

# 11453 ATels with GPT 4o (SORRY!)

- Query Simbad (to get MAIN\_ID):
  - GPT4o sources
  - REGEX sources
  - AstroBERT (open source LLM) sources
- 3024 ATels: the GPT4o sources == REGEX + AstroBERT sources
- 79 ATels: GPT4o found sources and REGEX + AstroBERT not
- 14 ATels: GPT4o not found sources, but REGEX + AstroBERT yes
- 8178 ATels: REGEX + BERT found additional sources compared to GPT4o
- 1589 ATels: GPT4o found additional sources compared to REGEX+BERT
- Sometimes GPT4o detects phrases like “2MASS source of mag 14”

# ATel #11007

```
"Celestial Object": [
  {
    "Name": "SN2017ili",
    "RA": "23:44:49.30",
    "Dec": "-12:43:58.80",
    "redshift": "0.025",
    "Type": "Ia",
    "Time": "Null",
    "Date": "20171117"
  },
  {
    "Name": "SN2017iln",
    "RA": "03:35:24.45",
    "Dec": "06:05:33.20",
    "redshift": "0.078",
    "Type": "Ia",
    "Time": "Null",
    "Date": "20171117"
  },
  {
    "Name": "SN2017igy",
    "RA": "02:32:10.35",
    "Dec": "04:04:39.9",
    "redshift": "0.073",
    "Type": "Ia",
    "Time": "Null",
    "Date": "20171113"
  },
  {
    "Name": "SN2017igt",
    "RA": "02:13:23.90",
    "Dec": "03:23:35.58",
    "redshift": "0.029",
    "Type": "Ic",
    "Time": "Null",
    "Date": "20171117"
  },
  {
    "Name": "SN2017ihf",
    "RA": "02:43:40.60",
    "Dec": "-24:11:52.86",
    "redshift": "0.025",
    "Type": "Ib",
    "Time": "Null",
    "Date": "20171110"
  }
],
"References": [
  "Smartt et al. 2015, A&A, 579, 40",
  "Tonry et al. 2011, PASP, 123, 58",
  "Tonry et al. ATel #8680",
  "Chambers et al. 2016, arXiv:1612.05560",
  "Blondin & Tonry, 2007, ApJ, 666, 1024",
  "Harutyunyan et al., 2008, A&A, 488, 383"
],
"Event Type": "supernova classifications",
"Time": "Null",
"Date": "2017 Nov 24",
"Telescope": "ESO New Technology Telescope",
"Instrument": "EFOSC2 and Grism 13",
"Observatory": "La Silla"
```

ePESSTO, the extended Public ESO Spectroscopic Survey for Transient Objects (see Smartt et al. 2015, A&A, 579, 40 <http://www.pessto.org>), reports the following supernova classifications. Targets were supplied by the ATLAS survey, see Tonry et al. (2011, PASP, 123, 58) and Tonry et al. (ATel #8680) and Pan-STARRS Survey for Transients (see Chambers et al. 2016, arXiv:1612.05560, and <http://pswww.ifa.hawaii.edu>). All observations were performed on the ESO New Technology Telescope at La Silla on 2017 Nov 24, using EFOSC2 and Grism 13 (3985-9315Å, 18Å resolution). Classifications were done with SNID (Blondin & Tonry, 2007, ApJ, 666, 1024) and GELATO (Harutyunyan et al., 2008, A&A, 488, 383). Classification spectra and additional details can be obtained from <http://www.pessto.org> (via WISeREP) and the IAU Transient Name Server.

Survey Name	IAU Name	RA (J2000)	Dec (J2000)	Disc. Date	Source	Disc Mag	z	Type	Phase	Notes
ATLAS17nja	<a href="#">SN2017ili</a>	23:44:49.30	-12:43:58.80	20171117	ATLAS	19.862	0.025	Ia	pre-max	(1)
ATLAS17njj	<a href="#">SN2017iln</a>	03:35:24.45	06:05:33.20	20171117	ATLAS	19.518	0.078	Ia	pre-max	(1)
PS17fcj	<a href="#">SN2017igy</a>	02:32:10.35	04:04:39.9	20171113	PS1	19.25	0.073	Ia	+3d	(1)
PS17fcg	<a href="#">SN2017igt</a>	02:13:23.90	03:23:35.58	20171117	PS1	19.68	0.029	Ic	near-max	(1)
ATLAS17ngb	<a href="#">SN2017ihf</a>	02:43:40.60	-24:11:52.86	20171110	ATLAS	18.547	0.025	Ib	+5d	(1)

(1) Redshift estimated with SNID or Gelato

# ATel #10360

```
"Celestial Object": [
{
  "Name": "Gaia17bdj",
  "RA": null,
  "Dec": null,
  "redshift": null,
  "Type": "candidate CV",
  "Time": "16:41",
  "Date": "2017 May 03"
},
{
  "Name": "Gaia17bdk",
  "RA": null,
  "Dec": null,
  "redshift": null,
  "Type": "candidate microlensing event",
  "Time": "16:53",
  "Date": "2017 May 03"
},
{
  "Name": "Gaia17bdm",
  "RA": null,
  "Dec": null,
  "redshift": null,
  "Type": "candidate CV",
  "Time": "16:34",
  "Date": "2017 May 03"
}
],
"References": "ESA Gaia, DPAC, Photometric
Science Alerts Team
(https://gaia.ac.uk/selected-gaia-science-alerts)",
"Event Type": "Astronomical observations",
"Time": "8-9 UT and 4-5 UT",
"Date": "2017 May 03 and 2017 May 05",
"Telescope": "Swiss 1.2m Euler telescope",
"Instrument": "ECAM",
"Observatory": "La Silla Observatory, Chile"
```

We report confirmation of [Gaia Science Alerts](#) transients [Gaia17bdj](#), [Gaia17bdk](#), and [Gaia17bdm](#). Images were obtained using modified Gunn R-band and Geneva V-band filters on the ECAM instrument installed on the Swiss 1.2m Euler telescope at La Silla Observatory, Chile, on 2017 May 03 between 8 and 9 UT, and on 2017 May 05 between 4 and 5 UT.

[Gaia15bdj](#) is a candidate CV (brightened by 4 mag) with alternative ID ASASSN-15lc and is seen as a faint source in an archival DSS image.

[Gaia15bdk](#) is a candidate microlensing event in the Galactic bulge that has brightened by nearly 2 magnitudes and is now rapidly declining in brightness.

[Gaia15bdk](#) is a candidate CV with no host apparent in archival DSS images and alternative ID SSS120620:1751-605942.

400x400 pixel cut out images in ECAM modified Gunn R band are available via the links above (N is up, E is left and the scale is 0.215"/pix). The below table provides heliocentric Julian Date - 2,400,000d and magnitudes.

57876.84 -->

Source	HJD - 2400000	R- band [mag]	V- band [mag]
Gaia15bdj	57876.85	16.69	16.57
Gaia15bdj	57878.70	17.65	--
Gaia17bdk	57876.86	16.58	--
Gaia17bdk	57876.83	18.58	18.40
Gaia17bdk	57878.69	18.92	--

The photometric calibrations was performed using the [Cambridge Photometric Calibration Server \(CPCS\)](#), designed and maintained by Sergey Koposov and Lukasz Wyrzykowski. We acknowledge ESA Gaia, DPAC and the Photometric Science Alerts Team (<https://gaia.ac.uk/selected-gaia-science-alerts>).

# RA DEC

- GPT4o RA-DEC positions for 30345 sources:
  - for 15290 sources: GPT4 found positions & they are in the text
  - for 14079 sources: GPT4 found NONE position
    - we did not compare with REGEX and astroBERT, yet
    -
  - there are 977 sources with “WEIRD” situation:
    - e.g. extract position from the name: ATEL #5664 - Source: PSN J11500016+2120166
    - Multiple positions per source (need more studying... )
    - Multiple formats per sources

# Next steps

- Query TNS and Fink for better source check
- Find simple/fast ways to convert RA-DEC positions to a given standard
-