

Standardizing VOEvents and archives

Emily Petroff*†

Anton Pannekoek Institute, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands

E-mail: e.b.petroff@uva.nl

Dave Morris

Institute for Astronomy, School of Physics and Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ, UK

Eric Chassande-Mottin

APC, Univ Paris Diderot, CNRS/IN2P3, CEA/Irfu, Obs. de Paris, Sorbonne Paris Cité, F-75013 Paris, France

The aim of this workshop was to discuss possibilities for standardizing VOEvent messages and infrastructure across the transient community. VOEvents are currently used to communicate information about a wide variety of astrophysical transients, thus a total standardization of the information contained within a VOEvent message may not be possible. However, this workshop focused on changes that could be made to VOEvent entry fields, archiving, and querying that would make the VOEvent framework more user friendly and future proof.

The New Era of Multi-Messenger Astrophysics - Asterics2019

25 - 29 March, 2019

Groningen, The Netherlands

*Speaker.

†Veni Fellow.

1. Introduction

An ever increasing number of astronomical transients are discovered yearly in data from observatories around the world. Communicating information about these new detections quickly and efficiently is essential to facilitating multi-wavelength and multi-messenger transient science. Many transient discoveries are now transmitted publicly over the Virtual Observatory (VO) network through the use of `xml` messages called VOEvents¹. VOEvents offer several advantages in facilitating transient science: they are machine readable, can be generated by an open-source suite of software tools², contain only the essential observables related to a detection, and are sub-divided into several elements that describe things like event time, sky location, and physical properties.

The VO network and the VOEvent framework are developed and maintained by the International Virtual Observatory Alliance (IVOA). A number of tools and libraries have been developed using the IVOA standard to interface with the network and facilitate use by the community such as the Comet broker system³ [1] and the 4 Pi Sky transient research project⁴ [2]. These systems are currently used by many transient facilities, such as the Neil Gehrels *Swift* satellite which communicates new gamma ray burst (GRB) detections via VOEvent which can be subsequently followed up at other facilities [3].

While VOEvents have been put to good use in the transient community, the focus of this workshop was to discuss possible improvements to the framework. In particular, this workshop focused on the topics of standardizing VOEvent messages, making VOEvents compatible with the FAIR data principles, facilitating easier querying of past events, and adding new fields to VOEvents to make them more useful for new types of transients. We briefly discuss the outcomes of these discussions and the resulting recommendations in the following sections.

2. Standardization of VOEvent reporting

One of the distinct advantages of the VO system is that VOEvents are incredibly flexible in terms of the properties that are reported in an individual message. VOEvents themselves are structured with various elements (Who, What, WhereWhen, Why, and How), with the `What` element containing all the relevant observed parameters of the event. While this flexibility is a strength of the system, the lack of standardization in what is reported in a VOEvent message can cause confusion or difficulty in parsing incoming messages by facilities wanting to participate in follow-up.

The fast radio burst (FRB) community has attempted to solve this problem by creating a standard message framework with defined names for all parameters relevant to a new FRB [4]. In this case, a standard set of parameters describing the event are required (e.g. dispersion measure, event duration, signal-to-noise ratio, peak flux density) and all others are optional. The effort was also motivated by the use of VOEvents to populate a centralized fast radio burst catalogue [5] which can import new events directly from incoming messages⁵.

¹<https://wiki.ivoa.net/twiki/bin/view/IVOA/IvoaVOEvent>

²<https://voevent-parse.readthedocs.io/en/stable/index.html>

³<http://comet.transientskp.org/en/stable/>

⁴<https://4pisky.org/>

⁵<https://www.frbcatalog.org>

In this workshop, the participants discussed developing a similar standardization for other types of transients such as short and long GRBs, X-ray flares, supernovae, etc., with parameters relevant to each type of transient. While such an effort was possible in the FRB community, this was undertaken when the community was still relatively small and the standard was developed before any FRBs had been announced over the VO Network. This approach may be challenging in larger and more established transient fields that already have some systems in place for generating and receiving VOEvents. While an attempt at standardization would require the agreement of a larger community of users in other sub-fields of transient astronomy, it was recommended in this workshop that smaller groups within each community draft a proposed standard to be commented on and modified by the larger user base.

- *Recommendation: Smaller working groups within transient sub-fields develop a proposed standard to be iterated upon within the wider user base. Once a standard is developed, it can be submitted to the IVOA and published as an IVOA Note.*

3. FAIR Data principles

The FAIR Data principles have been developed to better facilitate reuse of scientific data and focus on four key principles to which data should adhere [6]. Specifically data should be **F**indable, **A**ccessible, **I**nteroperable, and **R**eusable. In this workshop we examined whether VOEvents currently meet these criteria and how they could be improved to better adhere to FAIR principles.

The criterion on which VOEvents perform best is their interoperability. VOEvents already have standardized metadata and structure that adheres to a single system. VOEvents also perform fairly well in terms of accessibility; the `What` element in a VOEvent can contain a wealth of information on an event and the VOEvent metadata is understandable to humans and machines. However, there are some issues related to accessibility as well. For example, the `Why` element of a VOEvent contains an importance value related to the judged validity of probability of the transient; this scoring is in no way standardized, either within transient community or more broadly over all VOEvent users.

VOEvents perform more poorly on metrics related to reuse and findability. In terms of reuse, VOEvents often lack information on provenance (i.e. how was this event judged to be a ‘fast radio burst’) and terms of reuse such as a license or citation guideline can be unclear. The biggest issue, however, is related to the findability of VOEvents. At present, no robust system exists for querying past events and finding VOEvents in, i.e. a particular region of the sky, date range, or waveband. The specific issue of querying past events is addressed in the next section, but some discussion time in this workshop was given to making FAIR data principles a priority as we begin to develop new capabilities for VOEvents. Improving descriptions of the algorithms and criteria used for transient identification as well as providing external links to appropriate papers/pipelines for citation would already provide a big step towards FAIR compatibility.

- *Recommendation: All future updates to the VOEvent framework should adhere to FAIR data principles.*
- *Recommendation: Authors of new events should provide URLs to external pages describing event provenance as well as citation or license information.*

4. Querying past VOEvents

In the context of the FAIR data principles discussed above, one topic of discussion at this workshop was the need for better tools to query past VOEvents including more organized archiving tools. A robust querying infrastructure would make VOEvents much more findable. Ideally, when a new transient event is found, you would like to know if other facilities detected a transient in the error circle of your detection either at the same time or in the past, if other facilities were observing the field at the same time, and if other observatories have looked at this field previously. Tools for querying should also take into account the observation error box, both of the new transient event and previous transients reported nearby.

In this context, a successful system would require that facilities report their observing plans to be queried later. It was noted in this discussion that some observatories may be unwilling to report their observing plans ahead of time, even if scheduled far in advance, and may instead wish to publish an archive of past schedules after the fact. While the latter case may be less conducive to real-time science, any commitment from observatories to make their pointing schedules public would be incredibly useful.

- *Recommendation: Observatories should publish schedule or metadata archives of where they have observed and at what time to query against new VOEvents.*

5. Requested changes to VOEvents

In addition to the discussion points raised above, the workshop organizers also asked the participants for feedback on what aspects of the VOEvent structure should be changed or if new parameters should be added. The following recommendations were made by the community:

- *Inclusion of non-standard error regions to describe irregular beam shapes, or multiple non-contiguous sky regions, e.g., through the use of HEALPix Multi-Order Coverage localization skymaps⁶.*
- *A new entry in the *Why* element for ‘source type’ with standard entry options, e.g. FRB, sGRB, LGRB, LMXB, ...*
- *Some definition or flag in the VOEvent metadata to identify the name of the service and/or stream of transient alerts from which the event was generated.*

References

- [1] J. Swinbank, *Comet: A VOEvent broker*, *Astronomy and Computing* **7** (2014) 12 [1409.4805].
- [2] T. D. Staley and R. Fender, *The 4 Pi Sky Transient Alerts Hub*, *arXiv e-prints* (2016) arXiv:1606.03735 [1606.03735].
- [3] G. E. Anderson, T. D. Staley, A. J. van der Horst, R. P. Fender, A. Rowlinson, K. P. Mooley et al., *The Arcminute Microkelvin Imager catalogue of gamma-ray burst afterglows at 15.7 GHz*, *MNRAS* **473** (2018) 1512.
- [4] E. Petroff, L. Houben, K. Bannister, S. Burke-Spolaor, J. Cordes, H. Falcke et al., *VOEvent Standard for Fast Radio Bursts*, *arXiv e-prints* (2017) arXiv:1710.08155 [1710.08155].

⁶<https://emfollow.docs.ligo.org/userguide/>

- [5] E. Petroff, E. D. Barr, A. Jameson, E. F. Keane, M. Bailes, M. Kramer et al., *FRBCAT: The Fast Radio Burst Catalogue*, *PASA* **33** (2016) e045 [1601.03547].
- [6] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak et al., *The FAIR Guiding Principles for scientific data management and stewardship*, *Scientific Data* **3** (2016) 160018.