Challenges of Trustable AI and Added-Value on Health
B. Séroussi et al. (Eds.)

doi:10.3233/SHTI220480

A Conceptual Framework for Representing Events Under Public Health Surveillance

Anya OKHMATOVSKAIA[a,1] Yannan SHEN[a], Iris GANSER[a,b], Nigel COLLIER[c], Nicholas B. KING[a], Zaiqiao MENG[c] and David L. BUCKERIDGE[a]
[a] *School of Population and Global Health, McGill University, Canada*
[b] *Bordeaux Population Health Research Center, Bordeaux University, France*
[c] *Department of Theoretical & Applied Linguistics, University of Cambridge, UK*

Abstract. Information integration across multiple event-based surveillance (EBS) systems has been shown to improve global disease surveillance in experimental settings. In practice, however, integration does not occur due to the lack of a common conceptual framework for encoding data within EBS systems. We aim to address this gap by proposing a candidate conceptual framework for representing events and related concepts in the domain of public health surveillance.

Keywords. Event-based surveillance, conceptual modelling, outbreak detection

1. Introduction

Event-based surveillance (EBS) is the process of monitoring for, and reporting on, events of potential public health importance using information sources not specifically designed for this purpose, such as online news articles, social media, drug sales, or absenteeism data [1]. Compared to indicator-based surveillance, which relies on confirmed cases of known diseases, EBS is characterized by potentially higher sensitivity and timeliness, but usually at the cost of specificity. EBS systems face unique challenges, including questionable source reliability, low signal-to-noise ratio, and biases introduced by uneven internet coverage and trends in media reporting. Due to differences in scope, orientation, and design, the performance varies significantly between EBS systems and research has shown the benefit of integrating information across multiple systems [2]. However, for such integration to occur routinely in practice, semantic interoperability is necessary, i.e., the representation of key concepts must be aligned among systems.

At the core of all EBS systems is the concept of an *'event'*. While this concept appears intuitively clear, its use is inconsistent across different systems, particularly in relation to the terms *'signal'* and *'outbreak'*. A common approach is to view an event as a stage in the alerting process, for example: "Once the signal has been triaged and verified, it becomes an event" [3]. This narrative is ontologically problematic and assumes an implementation-specific event definition, thus making it impossible to compare different systems, as their decision processes vary. Another common narrative suggests that there are two different kinds of events: 1) an occurrence of public health significance (e.g. disease outbreak), and 2) other events (e.g. an increase in drug sales or Google searches)

[1] Corresponding Author, McGill University Clinical and Health Informatics, 2001 McGill College Avenue, Suite 1200, Montréal, Québec H3A 1G1, Canada; E-mail: anna.okhmatovskaia@mcgill.ca.

that may indicate or signal the occurrence of public health significance: "The goal of EBS is to detect unusual events that might signal an outbreak." [4]. However, there is no explicit model to describe how those other events relate to a disease outbreak.

Reaching consensus regarding these high-level concepts should create a foundation for achieving semantic interoperability amongst systems, improving global EBS. In this paper, we propose a conceptual framework for representing events and related concepts in the context of EBS that can serve as a foundation for addressing these issues. In this initial modelling stage, we do not assume a specific formal encoding of our framework.

2. Related work

Formally defining and representing the dynamic aspects of the world, such as events and processes, has been a subject of multiple disciplines from philosophy to linguistics to artificial intelligence. Yet the definitions of event vary across different frameworks, and numerous, often incompatible, concept hierarchies have been proposed [5].

2.1. Upper-level ontologies and foundation models of events

Events are positioned at the top level of the concept hierarchy and are typically defined as things that happen in time (perdurants, occurrents), as opposed to things that maintain their existence (endurants, continuants). Events are included in all major upper-level ontologies, but alternative terms are often used to refer to them, e.g. 'eventive perdurant' [6], 'transition' [7], 'process' [8]. Events usually imply change over time and are distinguished from 'qualities' and 'states', with the exception of some models that consider static states as a type of event [9]. Besides defining events, existing event frameworks typically discuss event typology, their spatiotemporal relations, mereology, participation of non-event entities, and causal relationships.

Most frameworks agree that events have temporal boundaries, i.e. a distinct start and end time points. Allen & Fergusson [10] point out the arbitrary nature of these points, while in basic formal ontology (BFO) [8] the definition requires them to "correspond to real discontinuities". The relation of events to places is sometimes modelled as direct [11], or indirect, via participating agents or objects [9, 12, 13].

Events can be comprised of other events as proper parts. Guizzardi et al. [12] suggest the existence of atomic events that cannot be further subdivided, but this view is not supported by other models, which point out that atomicity of events can only be assumed at a chosen level of temporal granularity. Some authors emphasize the importance of an observer in identifying event boundaries and parthood. According to [10], events do not really exist, but are the way by which agents classify the observed patterns of change.

Kaneiwa et al. [9] categorize events into types according to participating agents and objects (e.g. natural event vs. artificial event) and distinguish several semantic functions of events (state change, spatial existence change, etc.). Rich Event Ontology (REO) defines many subclasses of events, including life event, cognitive event, motion, etc. [14].

2.2. Domain applications of event models

Brown et al. have pointed out that despite the proliferation of software ontologies, most have focused on developing rich object hierarchies, but not event hierarchies [14]. Indeed, upper-level ontologies, e.g., [6, 8], only define the general aspects of events at the most abstract level, and few ontologies extend this concept. REO is one example of such an extension, providing a rich structure of events at varying levels of specificity across many domains, and relating events to their key objects and participants [14]. Other domain applications of formal event models include the Music Ontology [11], the Time Event Ontology for representing clinical events [15], and Adverse Event Ontology [16]. Among domain-specific event ontologies, BioCaster Event Ontology [17] is the most relevant to the work presented here: this ontology built with an natural language processing focus extends Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) to provide a multilingual vocabulary for extracting infectious disease outbreak information from media reports. Here, we build on this work to propose a conceptual event model with the aim of improving interoperability in public health surveillance.

3. Conceptual Framework Description

In the proposed conceptual framework, we intend to tease apart, define, and clarify the relationships among several phenomena, which are often described inconsistently and sometimes confused in the context of EBS: a) public health (PH) events that the systems are trying to detect (e.g. E. Coli outbreak at a restaurant due to tainted meat), b) individual cases of illness or clusters of cases that are part of a PH event, c) events that are causally related to an underlying PH event (e.g. school absences, increase in the sales of cold remedies), and d) documents about events (e.g. news articles, social media posts).

3.1. Definitions

Our definition of an *event* closely matches a definition of a 'process' in BFO, and is similar to 'eventive perdurant' in DOLCE. An *event* is an occurrent that exists in time, composed of temporal parts, and is a maximally connected spatiotemporal whole with bona fide beginnings and endings, corresponding to real discontinuities

A *public health event* is usually defined as an occurrence that involves disease or death above expected levels for a given time and place. Thus, a public health event can be formalized as a subtype of event that involves a *causative agent/hazard*, one or more *affected populations* as participants, and comprises *sub-events* (parts) that involve cases of illness or death. Although qualifying incidence of disease or death as being "above expected level" is central to the definition of a PH event, we do not formally capture this aspect because of its contextual and subjective nature.

We adopt a view that partitioning of public health events into sub-events depends on the context and the observer (e.g., a reporting news source). A sub-event can be an individual case of illness or a cluster of cases. Thus, some public health events are observed as a sequence of many sub-events, and some – as just one (e.g., a case of anthrax).

In addition to PH events and their sub-events, other subtypes of event can be relevant to representing and reasoning about PH events, when they are causally related to a PH event or any of its parts and can serve as signals for EBS; these are '*associated events*.'

3.2. Properties and relations

At a minimum, PH events are characterized by time, location, hazard type and affected population. Events unfold over a *time interval* with distinct and ordered *start time* and *end time*, and within certain geographic *location*, which are often unobservable directly.

PH events can be related to other events via mereological (part-of), temporal, and causal relations. Temporal relations between events of any granularity level are modelled in our framework using Allen's interval algebra. Recognizing the challenges of formally modelling causality, we propose to use a simple 'has-effect' relation for this purpose.

Various subcategories of PH event introduce additional properties. For example, infectious disease outbreaks are associated with a *disease* and a *pathogen*, from which other properties relevant to an outbreak can be inferred, such as the *mode of transmission*.

3.3. Event categories

The World Health Organization classifies PH events by hazard type into the following categories: zoonotic, chemical, food safety, infectious, natural disaster, nutritional deficiency, medical product, radio-nuclear, and undetermined [18]. Infectious disease outbreaks are the main focus of EBS systems, comprising 76%-93% of all events [18]. Within the class of infectious disease outbreaks, two broad types of events are noteworthy from a practical perspective: 1) relatively small communicable *disease clusters*, localized in time and space, and 2) *epidemics* (e.g., COVID-19) that extend over long time intervals and large geographic regions and have complex dynamics (changing geography, multiple waves). The distinction between these two types is not ontologically strict, but reflects two different workflows in EBS.

While almost any kind of event can become causally related to a PH event or its sub-events, based on our analysis of existing EBS systems we include in our framework the following categories of associated events as particularly relevant: environmental events, natural and artificial (e.g. increase in mosquito activity, chemical plant incident); health-related behaviors (e.g. drug purchases); health care system events (e.g. hospitalizations, lab tests); interventions (e.g. travel ban); and social events (e.g. panic, protests). Existing ontologies can be used to further classify events within each of these categories.

3.4. Event documentation

Following [13], we believe that it is important to represent documents (e.g. media reports) about public health events, their sub-events, and associated events explicitly. Documents are information objects that represent events, and as such, they have a set of properties, distinct from that of events. Documents typically contain only partial and, possibly, inaccurate or distorted information about an event, increasing the possibility of introducing false positives or negatives. Also, multiple documents can refer to the same event, complementing or contradicting each other. A wide-spread issue in EBS of conflating events and event reports, poses a barrier to obtaining accurate event information.

4. Discussion and Future Work

Formal event models, including some intended for use in public health surveillance, have been proposed in the past. However, these models have not been adopted widely, possibly because their development was motivated by issues within individual systems, such as improving the parsing of news reports. Building on this prior work, our framework is developed based on the analysis of conceptual similarities and discrepancies among several existing EBS systems, aiming at creating a harmonized vocabulary with a potential for broader adoption. The alignment with a widely used upper ontology, BFO, should facilitate further incorporation of relevant domain knowledge from numerous BFO-mapped ontologies for a variety of applications.

Although the development of this framework has benefited from our interactions with colleagues who work with EBS, an important next step is to engage with EBS system stakeholders to obtain feedback on the proposed model regarding its utility and potential for adoption. After consulting with stakeholders and revising the framework based on their feedback, we intend to proceed to encoding this framework as a formal ontology. We will also elaborate use cases for application of the ontology within and across existing EBS systems. In the longer term, we can see potential value in extending a typology of associated events and building a rich causal model to link associated events to PH events to support reasoning for outbreak detection.

References

[1] O'Shea J. Digital disease detection: A systematic review of event-based internet biosurveillance systems. Int J Med Inform. 2017 May;101:15-22.

[2] Barboza P, Vaillant L, Mawudeku A, Nelson NP, Hartley DM, Madoff LC, Linge JP, Collier N, Brownstein JS, Yangarber R, Astagneau P; Early Alerting Reporting Project Of The Global Health Security Initiative. Evaluation of epidemic intelligence systems integrated in the early alerting and reporting project for the detection of A/H5N1 influenza events. PLoS One. 2013;8(3):e57252.

[3] Balajee SA, Salyer SJ, Greene-Cramer B, Sadek M, Mounts AW. The practice of event-based surveillance: concept and methods. Global Security: Health, Science and Policy. 2021 Jan 1;6(1):1-9.

[4] CDC Division of Global Health Protection (DGHP). Event-based surveillance. https://www.cdc.gov/globalhealth/healthprotection/gddopscenter/how.html; cited Jan 26, 2022.

[5] Galton AP. The ontology of states, processes, and events. 2012

[6] Gangemi A, Guarino N, Masolo C, Oltramari A, Schneider L. Sweetening Ontologies with DOLCE. In: Gómez-Pérez A, Benjamins VR, editors. Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web: 13th International Conference, EKAW 2002. Proceedings. Berlin, Heidelberg: Springer; 2002. p. 166--181.

[7] Pustejovsky J. The syntax of event structure. cognition. 1991 Dec 1;41(1-3):47-81.

[8] Smith B, Kumar A, Bittner T. Basic formal ontology for bioinformatics. 2005

[9] Kaneiwa K, Iwazume M, Fukuda K. An upper ontology for event classifications and relations. InAustralasian Joint Conference on Artificial Intelligence 2007 Dec 2 (pp. 394-403). Springer, Berlin, Heidelberg.

[10] Allen JF, Ferguson G. Actions and events in interval temporal logic. Journal of logic and computation. 1994 Oct;4(5):531-79.

[11] Raimond Y, Abdallah SA, Sandler MB, Giasson F. The Music Ontology. InISMIR 2007 Sep 23 (Vol. 2007, p. 8th).

[12] Guizzardi G, Wagner G, de Almeida Falbo R, Guizzardi RS, Almeida JP. Towards ontological foundations for the conceptual modeling of events. International Conference on Conceptual Modeling 2013 Nov 11 (pp. 327-341). Springer, Berlin, Heidelberg.

[13] Scherp A, Franz T, Saathoff C, Staab S. A core ontology on events for representing occurrences in the real world. Multimedia Tools and Applications. 2012 May;58(2):293-331.

[14] Brown SW, Bonial C, Obrst L, Palmer M. The rich event ontology. In Proceedings of the Events and Stories in the News Workshop 2017 Aug (pp. 87-97).

[15] Li F, Du J, He Y, Song HY, Madkour M, Rao G, Xiang Y, Luo Y, Chen HW, Liu S, Wang L, Liu H, Xu H, Tao C. Time event ontology (TEO): to support semantic representation and reasoning of complex temporal relations of clinical events. J Am Med Inform Assoc. 2020 Jul 1;27(7):1046-1056.

[16] Ceusters W, Capolupo M, De Moor G, Devlies J, Smith B. An evolutionary approach to realism-based adverse event representations. Methods of information in medicine. 2011;50(01):62-73.

[17] Kawazoe A, Chanlekha H, Shigematsu M, Collier N. Structuring an event ontology for disease outbreak detection. InBMC bioinformatics 2008 Apr (Vol. 9, No. 3, pp. 1-8). BioMed Central.

[18] World Health Organization. Acute public health events assessed by WHO Regional Offices for Africa, the Americas and Europe under the International Health Regulations (2005) 2016 Report.