Exploring the Ontology of Pandemic

C. Maria Keet

Department of Computer Science, University of Cape Town, South Africa

Abstract

Pandemics do take place. When exactly they begin and end, and why, is harder to determine, as also demonstrated in early 2020 at the start of the Covid pandemic and the many debates in 2022 on calling it over. To determine these points, one has to know which criteria have to be satisfied and which not, respectively. This requires a clear definition of what a pandemic is, with at least its necessary and sufficient characteristics. There is no such crisp and clear definition, neither in the expert documentation nor in domain ontologies. In this paper, we assess mentions of 'pandemic' in domain ontologies, evaluate the argument that foundational ontologies may provide guidance, and examine the characteristics that domain experts have put forward for pandemics. The guidance from foundational ontologies is underwhelming when taken together, but tooling greatly simplified the alignment. The assessment of characteristics show that pandemic is not bearer of them all but they are of attendant entities, elucidates which ones are dependent and which essential, and it demonstrates why one may compute more than one unique start and end of a pandemic. Considering the complexities, it may be of use to develop an ontology of pandemics.

Keywords

Pandemic, Foundational Ontology, Ontology comparison, Covid-19

1. Introduction

Pandemics have taken place throughout the millennia [1] and they have been investigated by scientists in many disciplines, but not ontology. Pandemics somehow start and end; e.g., the 1918 influenza pandemic ended in 1920. This raises the questions as to what the criteria are such that something is an instance of Pandemic and which criteria would have to be *not* met to determine the end of the pandemic. And, practically at present: how does it apply to the current Covid-19 pandemic?

They are simple questions without easy answers. The informal short description of 'pandemic' is that it is a large epidemic [2] and, by the WHO's guidelines, the pandemic is over when it is alike a seasonal influenza epidemic ("post-pandemic phase 6")[1]. While the former is easy to communicate to the public and the latter is operationally clear, scientifically, it is an unsatisfying answer. The key issues are that, ontologically, there is no clear definition of pandemic in such a way as to unambiguously classify something as being a pandemic or not being a pandemic. This holds for both the scientific literature of the subject domain and, as we shall see, the domain ontologies that have 'pandemic' in their vocabulary. In this unclear situation, foundational ontologies (FOs) should be able to be of assistance somehow, to help determine at least the category that a domain entity such as 'pandemic' is. There is only limited published independent practical alignment guidance for two of the multitude of the FOs [3, 4], in the sense of not having to rely on human services [5] to accomplish the task.

In taking steps toward the ontology of pandemic, we are guided by the following questions. What is a pandemic, ontologically? What properties does it have? To which entity in a FO should it be aligned or categorised into, and as part of that: how to align and how to know one has done that alignment correctly? Methodologically, we first examine related work with the 9 relevant domain ontologies and summarise the key advances from the domain experts. We then proceed toward an ontological characterisation, by, first, assessing 7 FOs and seek to align pandemic to it, and subsequently we examine the asserted characteristics of pandemics (notably [2], but also [6]) as an ongoing process of modelling refinements.

The tooling support for FO alignment was found to be very helpful and increased confidence that the alignment is as precise as it can be. The assessment of asserted characteristics of pandemics are, in some cases, actually properties of intricately related entities (e.g., the disease), some turned out to be dependent or implied, others are essential. Crucial for classifying something as (not) a pandemic, is that several properties are vague and thus there can be multiple start and end points for a single pandemic. To be able to classify things as a pandemic, a separate ontology of pandemics may be needed and perhaps a rule-based decision table as well.

In the remainder of the paper, the related work is presented in Section 2 and the assessment with FOs (Section 3) and pandemic's claimed characteristics (Section 4) follow afterward. We discuss in Section 5 and conclude in Section 6.

ICBO 2022, September 25-28 Ann Arbour, MI, USA
mkeet@cs.uct.ac.za (C. M. Keet)
0000-0002-8281-0853 (C. M. Keet)

[1] https://www.who.int/influenza/resources/documents/pandemic_phase_descriptions_and_actions.pdf

Table 1
Summary of where 'pandemic' is positioned in the ontology that has it (VIDO and CIDO import the IDO; hence, pandemic is also a process there), with ID where applicable. Ontology references: prefix https://bioportal.bioontology.org/ontologies/ and then their abbreviation as listed.

Ontology or model	Pandemic wrt Epidemic	Parent type in the subject domain	Top-level Category
AURA	is a	infection	event
CRISP	the same	medicine (4000-0244)	–
GSSO	is a	–	event (EFO_0009629)
IDO	sibling	–	process (*sensu* BFO, so not an event) (BFO_0000015)
IOBC	sibling	infection (200906067292879338)	'Biological phenomenon, process, and state' (System_cat)
NCIT	is a	'disease or disorder' (C2991)	–
MESH	is a	'public health' (D011634)	–

2. Related work

We consider first the domain ontologies that contain 'pandemic' and subsequently summarise domain literature that, overall, have it ill-defined as well.

2.1. Domain ontologies

Domain ontologies that have something to do with pandemics were collected by means of a BioPortal search using the search string "pandemic" (without quotation marks), which returned 17 hits[2]. Of those, 9 refer to the entity (of which 3 with the same IDO entry), 1 is a list rather than an ontology (the ELD) once it was added as a top category afterthought (in LOINC), once as instance for just the pandemic in 1918 (in OMIT), and once the organism (in SNOMED CT), rather, and 4 hits are related things, such as 'Product delay due to pandemic' in MEDDRA. The 9 were analysed further and their respective key points are summarised in Table 1. We shall discuss each in turn, on i) how pandemic relates to epidemic (and outbreak), ii) what its parent is within the subject domain scope, and iii) what its domain-independent ancestor is within the realm of top-level entities typically in FOs.

Regarding how pandemic relates to epidemic, a key difference across the 9 ontologies is to have it either as subsumption, the same, or sibling, which cannot be all correct. The subsumption rationale is based on the notion that a pandemic is a *large* epidemic, having spread over more or larger regions and affecting more people, i.e., that that is an extra feature meriting the subsumption.

The sibling approach of the IDO and its related ontologies is based on the premise that a pandemic is an aggregate or collection of "multiple infectious disease epidemics", as indicated by the natural language definitions in the ontology (it lacks an axiomatisation). There are three problems with this approach. The first, and key, problem is the 'collection' of epidemics: collections (i.e., collectives, aggregates) of individuals, be they processes or objects, are categorically different kind of entities from single individuals and from the individuals that are member of it—both ontologically (e.g., [7]) and in FOs such as BFO with its object aggregate (and, in BFO v1, process aggregate)—but epidemic and pandemic are not categorically different kinds of things. There is a fiat boundary between an epidemic evolving into becoming a pandemic and then subsiding into one or more epidemics, as the recent IDO documentation also indicates (see figure 4 in [8]). A secondary issue that the IDO approach faces in particular, is how to determine the boundary of one epidemic from another (of the same disease with same causative agent) to be able to construct a collective, if it indeed is one, and, more fundamentally, what the respective identities of those co-occurring epidemics are. This is unclear; among others: is it one epidemic in two places that it jumped to—e.g., from Italy to South Africa—or do they count as two? How to count two purportedly separate ones when they touch: do they merge to become one large one or are they presumed to be overlapping distinct entities (that may not be identifiable as such)? A third issue, and minor for the current scope, is the definition for epidemic in the ontology's annotation field, mentioning "statistically significant increase in the infectious disease incidence" as determiner, but actually it is based on a threshold that is based on a moving epidemic method that relies on historic data for the country [9], for influenza at least. If epidemic and pandemic are siblings rather than the former subsuming the latter, a better argument has to be put forward. At this stage of the analysis, the argument for subsumption is more plausible.

Regarding pandemic's parent in the subject domain of the selected ontologies (see Table 1), it cannot be either of them. In particular, public health in MeSH and medicine in CRISP are merely intended as broad groupings of context, not ontologically a superclass, despite that the ontology formally asserts the latter. The infec-

[2] https://bioportal.bioontology.org/; last checked d.d. 2-3-2022

tion in AURA and IOBC and disease or disorder in NCI are all more specific as superclass than the previous two, but are attendant topics, rather. A particular infection, an instance of a disease (or disease course in an organism) or disorder each are a single thing, as an intrinsic whole, in an individual organism, whereas from outbreaks onward to pandemic, there are multiple related infections in multiple organisms in at least one region. Pandemics need infections to happen, but that does not make them infections. Similarly, there is a disease that causes a pandemic, but that does not make the pandemic a disease.

Lastly, the top-level category, for those that had one. The IOBC category 'Biological phenomenon, process, and state' is too imprecise for ontological analysis, although it does provide a general indication that pandemic is something that is happening, or occurring or perduring, in ontological terms. For the others, the key distinction between process and event—however it may be formalised in one's preferred logic—is that processes, at least in theory, can go on forever whereas events necessarily have both start and an end, i.e., they have a limited duration by definition whereas processes do not. We know from history that pandemics indeed do have an end, even though it may not be trivial to determine and have fiat boundaries.

In sum, the domain ontologies have been mostly helpful in indicating what a pandemic is not, and a possible avenue for the direction for FO alignment.

2.2. Domain literature

Domain experts have had multiple debates about defining what a pandemic is and they assessed it from multiple angles, including infectious diseases, immunology, epidemiology, and health policy. In the literature, most debate occurred from around 2009 during the H1N1 influenza pandemic ('swine flu') to about 2011.

WHO's eventual *pandemic phases* document consists of descriptions of the phases, but steers clear of a definition of pandemic (see fn. 1). Its first phase is where an animal influenza circulates but is not reported to jump over to humans and it goes up to phase 6 (where we still are with Covid-19 at the time of writing), where there are sustained community outbreaks in two or more countries in one region and at least one other country in another WHO region. It also has a "post-peak" phase that is still pandemic but the worst is deemed to be over, and a "post-pandemic" phase when there are "levels seen for seasonal influenza in most countries with adequate surveillance", revealing also that these phases were specified either with the expectation that the next pandemic would be an influenza pandemic or on the basis that seasonal influenza is acceptable loss. Seasonal influenza surveillance is well-established and, knowing that those values vary by country for a range of reasons [10], each country regularly calculates it by comparing it to their own historical data over the preceding years [9][3]. Informally, an epidemic is an occurrence where there are multiple instances of an infectious disease in organisms, for a limited duration of time, and that affects a community of said organisms living in a region, and the number organisms it affects exceeds the agreed-upon threshold compared to average historical data for that region.

Also within the WHO there used to be a lot of debate about defining and describing a pandemic [11, 6, 2], and several national centres of (infectious) diseases made attempts as well. The lowest common denominator, and oft-repeated phrase, is still that it is a large epidemic. Minor additions or refinements of features include that it is spread over the world, or at least multiple regions and continents, and that it usually affects many people[4], that it has to involve a new disease[5], and that there are out-of-season infections [6]. Morens, Folkers and Fauci's literature review on pandemics [2] resulted in a list of eight characteristics, which are: wide geographic extension, disease movement, high attack rates and explosiveness; minimal population immunity, novelty, infectiousness, contagiousness, and, with some caution since it is not often added explicitly, also severity. We shall analyse them further in Section 4.

Overall, there is a tendency to oversimplify the definition of a pandemic for the wider population—as a large epidemic—and leave it for the experts to debate the features, which features really count most, and at least for some of them, what a threshold would be for declaring something an outbreak or an epidemic, and to keep it that way to avoid public confusion [2]. This approach of outward underspecification seems to have carried over to the domain ontologies. While an underspecification may suffice for some tasks, e.g., literature annotation, to be able to determine whether we have a pandemic at hand, when a pandemic can be called over, and to compare pandemics, a higher level of precision is required.

3. Categorising Pandemic

Before assessing the possible characteristic properties of a pandemic, it first has to be established what sort of generic entity it is. Two FOs have an assistive method to categorise an entity, therewith contributing to filling the tool gap in the ontology-as-a-service approach [5], which we therefore commence with. Afterward, we will

[3] The interested reader may consult such a surveillance report for an impression, e.g., https://www.nicd.ac.za/wp-content/uploads/2022/02/NICD-Private-Consultations-Surveillance-Epidemic-Threshold-Report-Week-5-2022.pdf

[4] https://www.cdc.gov/csels/dsepd/ss1978/lesson1/section11.html

[5] https://www.healthdirect.gov.au/what-is-a-pandemic

consider a selection of other FOs to examine whether it will shed any further light on the matter.

3.1. FOs with decision diagrams: DOLCE and BFO

There is a "D3" decision diagram to guide entity alignment [3] to DOLCE [12]. The question trail is as follows:

- *Is [pandemic] something that is happening or occurring?* Yes (perdurant).
- *Are you able to be present or participate in [a pandemic]?* Yes (event).
- *Is [a pandemic] atomic, i.e., has no subdivisions of it and has a definite end point?* No (accomplishment).

From the viewpoint of the affected population, it may be unpleasant to state that pandemic is an accomplishment, but it is one from the perspective of the infectious agent. Regardless, it confirms that a pandemic unfolds in time and is a temporal entity with a limited lifespan. It will cease to be a pandemic at some point in time and evolve back to epidemic[6]. This does not help determining when it starts or ends, just that it does.

We developed the BFO Classifier[7] [4] for BFO v2 [14], as a tool wrapping around the new decision diagram. Pandemic is a process, as follows:

- *Does [pandemic] persist in time or unfold in time?* Unfold in time (occurrent).
- *What is [pandemic]'s relation to the universal spatiotemporal region?* Inhabits (occurrent, still)
- *Does [pandemic] have a proper temporal part?* Yes (process)
- *Is [pandemic] the sum of the totality of processes in a material entity's spatiotemporal region?* No (process, still)
- *Is [pandemic] a collection of disjoint part-processes?* None of the above (remains a process)

Regarding the last question, which is optional since 'process' may be a leaf category (or: its subclasses are not exhaustive): a pandemic is not a collection of disjoint part-processes if the part-processes all have to be instances of different types of processes. BFO's process need not have an end, however, whereas pandemics do. So one may argue that bfo:process is not as precise as could be for a general high-level category. For now, most relevant is that pandemic is in the occurrent branch of BFO, which is in agreement with DOLCE since dolce:accomplishment is a perdurant and bfo:occurrent aligns to dolce:perdurant [15].

3.2. Other FOs

A recent inclusive survey lists 37 resources as FOs and candidate top-level categorisations [16]. We created a selection based on the following considerations: commonly considered foundational or top-level, ample documentation, ideally there is also a file to inspect with the formalisation, diverse in ontological commitments, and the final selection represents a wide geographic distribution just in case that matters. This reduced the selection to BORO, GFO, SUMO, UFO, and Yamato, which will be discussed in alphabetical order. The outcomes are summarised in Table 2.

BORO The Business Objects Reference Ontology [17] has the distinguishing metaphysical commitment of perdurantism. Since Type is the most specific for types, it does not reveal any category for Pandemic. BORO may be more useful for representing a pandemic's components or stages that have boundaries and parts, such as a wave of infections with a variant or a flare-up of infections due to lapsing prevention measures.

GFO The General Formal Ontology [18] differentiates between universals and individuals, where the Individual branch is where universals are placed for representing instances. By elimination of options, Pandemic has to be a Concrete and among those subclasses, an Occurrent that has temporal parts. Based on the descriptions in the ontology, in the documentation, and knowledge of pandemic, it is not possible to be certain of the appropriateness of Occurrent's subclasses, partially because 1) the descriptions are not easily accessible to casual users, in turn partially due to the incomplete characterisation in the OWL file and partially because the documentation is out of lockstep with the content of the ontology, and 2) Pandemic meets several criteria among the siblings. Notably, Actions "are occurents which are caused by some presential (the agent) at every (inner and outer) time-boundary of the chronoid framing the occurent.": there is an agent (the SARS-CoV-2 virus), so it may fit. And processes "are a special kind of occurrent. Processes are directly in time, they have characteristics which cannot be captured by a collection of time boundaries.", and where discrete processes "are made up of alternating sequences of extrinsic changes and states or continuous processes", which does not sound necessarily inapplicable either. Process and Action are not declared disjoint, so a multiple inheritance is permitted if needed. An argument in favour of action is that it has that agent for the duration of the action that the process does not have

[6] It might be that it evolves from pandemic to endemic, but even then it is expected to evolve via epidemic. Pandemic to endemic has not happened with any of the previous pandemics. HIV/AIDS is still categorised as a pandemic [13], not endemic. To not overcomplicate the analysis, any possible endemicity of an infectious disease after a pandemic is assumed to occur via an epidemic stage.

[7] https://bfo-classifier.github.io/

Table 2
Summary of what 'pandemic' would likely be aligned to (or categorise as) for a selection of foundational ontologies.

Foundational Ontology	Alignment (conservative)	Comment
BFO	Process	An occurrent. Lacks the specification of definite end
BORO	Type	Only few options available. Perdurantist foundation
DOLCE	Accomplishment	It is an event that is a perdurant
GFO	Occurrent	Likely one or more of its subtypes: action or process
SUMO	Process	Likely one or more of its subtypes: causing unhappiness and natural process
UFO	Situation??	Based on elimination and no observed conflicts; disjoint from endurant and event
YAMATO	ordinary event	A refinement to intrinsic or extrinsic accomplishment depends on more insight into pandemic

explicitly, and so it may capture pandemic somewhat more precisely.

The reference article for GFO [18] also includes event, which "is a right boundary of a process", rather, and thus inapplicable.

SUMO The Suggested Upper Merged Ontology [19] would have pandemic to be at least a process, which may be of the continuing variety and those with a sure end to it (i.e., events). Process has 10 direct subclasses and further subclasses, and it may satisfy more than one of them. For instance, the subclass causing unhappiness "Any Process whose result is that the patient of the process is unhappy.": if 'patient' includes the infected organisms, pandemics surely qualify, but Natural process as "A Process that take place in nature spontanously." fits as well. These sibling classes are not declared disjoint. Biological process sounds potentially applicable too, but it is a subclass of internal process that happens within a single object, which is not the case with a pandemic that happens to many organisms simultaneously. Regarding natural process's subclasses, there are electrical and mechanical processes and resonance, neither of which applies to pandemic, and so it remains at being a natural process as best fit among the options in the ontology.

UFO The Unified Foundational Ontology has several versions. We consider its OWL version gUFO and both the UFO-B for events [20] and the recent UFO overview [21]. These recent key sources are not in agreement with each other, which hampers alignment of pandemic to something; a selection follows.

The latest UFO reference paper [21] describes very little about UFO-B. Presumably, Pandemic would be a Perdurant and UFO-B slotted in there, but it cannot. "Perdurants are individuals that unfold in time accumulating temporal parts. They are manifestations of dispositions and only exist in the past." [21] where a disposition "inheresIn a unique ConcreteIndividual" [20]. That contradicts with gUFO:Situation that is not a manifestation of a disposition ("A gufo:ConcreteIndividual that is a particular configuration of a part of reality which can be understood as a whole and in which entities stand in relations.") and the gUFO does not have Perdurant. There is gufo:Event that is a "A gufo:ConcreteIndividual that 'occurs' or 'happens' in time. They may be instantaneous or long-running.", without the baggage of dispositions.

Regarding pandemic, then: 1) they clearly also exist in the now and not only in the past, and they will occur in the future, thus it cannot be a perdurant, 2) if indeed the "nature of events as manifestations of objects' dispositions" [20] holds, then it is not at all certain that pandemics would be events, since there is no inherent disposition in an infectious agent of 'pandemic-causing' to be manifesting itself. Perhaps a population of organisms, under the assumption it would be an individual, may have a disposition of, say, 'propensity to suffer from a pandemic' that is being manifested. On dispositions, two papers from 2016 by the main proposers of UFO are referenced, that in turn pass the bucket of clarification to a book by Molnar and Powers from 2006, at which time I categorised this line to be a dead end.

Two more directions were pursued. First, contrary to other FOs where Walk is an example of process or event, it is a Mode in UFO (which is the parent of Disposition), which is a Moment that is an Endurant [21]. Moment in the documentation is gufo:Aspect, and the match to the mode is gufo:IntrinsicMode that is described as "A gufo:IntrinsicAspect that is not measurable.", which does not rhyme with the Walk, and so this direction was abandoned as well. Second, pandemic as a Situation may be plausible, and the slight differences across the sources do not fundamentally contradict Situation in [20], cf gUFO:Situation, above (that is disjoint from Event and Endurant); it is mentioned in [21] but not used in any formalisation. What the disposition and atomic event are is to be determined. Having roughly checked the Situation axioms in [20], which are richer than in gUFO, it does not seem to contradict. There is no example in either [20] or gufo.owl as a way to further check understanding. In conclusion, if something had to be chosen for alignment of pandemic, it might be Situation in UFO-B.

YAMATO Lastly, the Yet Another More Advanced Top-level Ontology [22], with YAMATO20120714.owl. It is definitely an event, since it mentions the start and end and has unity and wholeness. Among that, it is an ordinary_event (cf. instant), but any more precise is difficult to determine at this stage, with its subclasses intrinsic_accomplishment and extrinsic_accomplishment. For the extrinsic one, there is a process and start and end, with as example 'a walk in the park'; the intrinsic one "is an accomplishment by itself alone", such as 'a conference, a game, or a trip'. If we were to know whether a pandemic always ends by itself or that that is superseded by humans deciding it is over, we would be able to say for certain which of the two it is. Either way, for sure pandemic is an ordinary_event in YAMATO.

In closing, without any aid to help aligning or categorising a subject domain entity to a FO, it is nontrivial to do so and that with less certainty, largely due to insufficient detail that may not be consistent across the sources. For the alignments themselves, they are split along a divide of either something that possibly can go on forever (process) where there is no entity in the ontology that has something that unfolds and with an end to it, and ones that do have such type of entity (event, accomplishment). Whether it is an intentional act of omission of the former group (at least: BFO, GFO, and SUMO), is unknown; their documentation does not indicate intentional exclusion.

4. Further domain analysis

So far, based on domain and foundational ontologies, pandemic is a subtype of epidemic, and it an occurrent (perdurant) that unfolds in time and, if available in the foundational ontology, an event or accomplishment. This is still at a very high level. In this section, we zoom in on what the properties of pandemic are claimed to be, what they may be ontologically, and to assess how, if at all, it may assist refining the ontological status of pandemic to determine when some thing is one or not.

The pandemic feature most often mentioned is the 'large'. Etymologically, the 'pan-' in pandemic means 'all', in that it is happening across the world or, practically, at least *multiple regions and continents*. For this to happen, it suggests that there must be an almost simultaneous transmission taking place, which entails that infections are happening *out of season*. Other proposed features include that it usually *affects many people* and that it has to involve a *new disease*. Morens, Folkers and Fauci collated eight characteristics of pandemics [2] and it is, to date, still the most comprehensive list. From an ontological and modelling viewpoint, 'characteristics' is to be understood informally, since there are multiple ways to represent them in an ontology, depending on the modelling style [23]. They are presented and discussed in the remainder of this section in the sequence they were presented in Morens et al.'s article.

1. Wide geographic extension This refers to the geographic region where the organisms that are infected with the disease-causing infectious agent are living. 'Wide' is vague, which may be fuzzy in the mathematical sense or be estimated by other means. It is the statement that indicates there is no crisp threshold when 'wide' starts (1/yes) or ends (0/no), but either with a gradual membership function between 0 and 1 or there are ranges of values based on some arbitrary scale. This could be precisiated by consensus at least in theory, but practically will still run into issues. For instance, 'if on 1 continent, then Localised', 'if 2 continents, then Wide', and 'if all or n-1 continents, then Global': not all continents are equal in size and population density (hence, disruptive effects of a growing epidemic), such rules are yet to be defined, and there are several different continent models, counting between 4 and 7 continents on earth. If there is scientific consensus on those aspects as well, then the vagueness can be eliminated.

For an ontology, three possible short-hand representations for the wide geographic extent are as a binary, say, geo_extent $\mapsto$ Pandemic $\times$ {local, wide, global}, where that range consists of either classes or values, or an object property geographic region and a suitable cardinality constraint, such as ≥ 4 for the 6 or 7-continent model. Alternatively, one may opt for including all components of the complex property—inclusive of the species, disease, and causative agent—for which we first need to address some of the other characteristics, further below.

Kelly's "out of season" feature [6] implies wide geographic extension and may be more useful than simply counting continents. Kelly focussed on influenza, which is seasonal, and many infectious diseases are seasonal [24], but it does not imply that *any* pandemic-causing infectious agent would otherwise be seasonal. A more accurate term may be *nonseasonal*; either way, this is a clear yes/no attribute of the disease. This is illustrated with the COVID-19 pandemic, where countries record infections throughout the year and have waves in multiple seasons.

2. Disease movement This means that there is transmission from one place to another place and it can be traced. This implies a comparison of geographic extents across time points, which has a known solution in information systems. Ontologically, at the type (TBox) level, however, it is computationally costly to represent and reason with temporal and spatial information. Depending on the purpose of the ontology, one could choose to represent those details nonetheless, e.g., using the DOL

Table 3
Summary of the properties of pandemics proposed by domain experts, whether it is an essential/necessary or mandatory property ("Req.?") or not, proxy terminology for somehow computed estimated values (in single quotation marks), the key bearer of that property, and any further comments.

Characteristic	Req.?	Value	Key bearer	Comment
Geographic extension	Y	'wide'	Disease	Prevalence of the disease
Disease movement	Y	yes	Disease	Measured over time and space
High attack rate	Y	'high'	Infectious agent	
Minimal population immunity	Y	'yes'	Population and immune system	Implies at least partial novelty
Novelty	Y	yes	Infectious agent and organism's immune system	Implies minimal population immunity
Infectiousness	Y	yes	Infectious agent	
Contagiousness	N	any type	Transmission mode	Implies infectiousness
Severity	N	'severe'	Disease and Pandemic	As compound property for both and computed in different ways
Out-of-season	N	yes	Disease	Implicit through wide geographic extension

framework to tie in a first order predicate logic [25], or to delegate that to an external system to merely record yes/no disease movement with a data property and a Boolean data type, or refine that into a categorical variable or with ranges how slow or fast it moves for more detailed analysis, where a 'yes' suffices for pandemic. Disease moment may not be an essential primary property of pandemics, but rather a secondary one related to arriving at, and maintaining, a wide geographic distribution.

3. High attack rates and explosiveness Many people are affected in a short amount of time. The notions of 'many', 'short', and 'high' all indicate vagueness, which can be made specific in different ways either with a membership function or thresholds by consensus. A shortcut with an attribute attackRate $\mapsto$ Pandemic $\times$ {low,medium,high} (or, noncommittal, to anyType) is useful for information systems, but not a scientific ontology. Attack rate relies on components to determine the rate, such as that there is a disease, an infectious agent, and the reproduction (R) number of the agent. The R number with all its variants, in turn, uses notions such as population, susceptibility to infection (or: level of immunity; see below), social dynamics of a population (i.e., how people behave), ability to measure infections and mortality figures, and a dispersion parameter (how (un)evenly an infected organisms passes on the agent to other individuals in the population)[8].

4. Minimal population immunity Immunity to an infectious agent is relative, in that an organism has it to a degree to some or all of the variants of the infectious agent, and likewise for the whole population. This is also an inherently fuzzy feature without crisp boundaries other than the two extremes of 100% immunity and 0% where no-one is, which is the case when the agent is so novel that not even partial cross-immunity—i.e., the immunity to another agent protects at least a little from the novel one—applies. To add this to a data model, one again can take the 'easy' way by introducing, e.g., a categorical variable alike a populationImmunity with possible values {minimal,partial,high,full}. Also this property can be unpacked to bring it more into the realm of ontologies. As a minimum, population, organism, and infectious agent are involved and, second, an agreed-upon definition of population immunity [27] is needed that, in turn relies on the R_0 estimate (and thus relates to the attackRate). A higher R_0 entails a higher population immunity threshold, as do imperfect vaccines when they do not prevent against infection [27]. While this latter aspect is less relevant for declaring a pandemic—minimal immunity is easy to determine—it is more relevant for when to declare a pandemic over, because it requires a certain adjustable level of immunity.

5. Novelty The species' immune system has never been exposed to it or, in the case of a localised epidemic, a subset of the population of a species has not been exposed to the infectious agent. This can be a yes/no attribute in its simplest form. One maybe could add 'partial' to make it three-valued, either in the sense of those subpopulations or a strain may be new but not the disease more broadly (as with influenza and SARS). This relates to immunity insofar as that novelty implies that there will be minimal population immunity. Entities underpinning novelty are that it is a property of the infectious agent in relation to the organisms of the species it infects, and therewith it is only secondary to pandemic itself.

[8] A brief overview about the R number and in the context of COVID-19 and the pandemic can be found in [26].

6. Infectiousness It has to involve an infectious disease and thus excludes non-infectious states, disorders, or diseases such as obesity, heart attacks, and smoking. Thus, there has to be an infectious agent that causes the disease, excluding all non-communicable diseases. This may be a straight-forward notion for which a 'yes' is essential to pandemics, alike infectiousness $\mapsto$ Pandemic $\times$ {yes} as a shortcut. Whose property it *really* is, is less straight-forward, however: a pandemic is not infectious, but the causative agent of the infectious disease that, in turn, gets a species into a pandemic, is the one that bears the infectiousness property, i.e., a *chain* of relations from, e.g., Pandemic $\sqsubseteq \exists$causedBy.InfectiousDisease to InfectiousDisease $\sqsubseteq \exists$causedBy.InfectiousAgent and only then InfectiousAgent $\equiv$ Agent $\sqcap \exists$infects.Organism, under the assumption of clear definitions of causation and infection and an agreement on whether the '$\exists$' over infects should really be a 'some' (or an 'only' or 'at most 1') or the converse.

7. Contagiousness This refers to the transmission process, i.e., how it is being spread, and thus entails the infectiousness property. Only a limited set of options of transmission are possible. For humans, it can be from person to person (i.e., contagious) or through some other medium, such as an animal intermediary (e.g., fleas, rats) or the environment (e.g., water, as with cholera), and among human-human, there are, touch, droplets, and airborne. While it is relevant to know from a scientific and health policy viewpoint which of the possible values it is, it is irrelevant for calling something a pandemic [1] or calling an end to a pandemic. That makes this an optional property for classifying something as a pandemic.

8. Severity While Morens et al. note that severity is not always included, they also observed that, historically, the term 'pandemic' has been applied more often for diseases that are severe or that have relatively high fatality rates, such as HIV/AIDS (a pandemic since 1981 [13]) and the three recent influenza pandemics of the past century, than for milder ones. As such, it would be only an optional property of the pandemic as a whole, or, more precisely, of the disease, not be part of the set of necessary and sufficient conditions. However, regarding assessing and representing this property, first, what is deemed severe and what not is subjective, and thus it will be either fuzzy or some thresholds can be set, alike mild, intermediate, severe, and very severe, or to WHO's [28] set of values for influenza ({below seasonal threshold, low, moderate, high, Extraordinary}).

Second, while severity is a property of the disease and whose value depends on, among others, number infections, case fatality rates, and treatment options, it is also a *compound property* of a pandemic. More precisely, e.g., WHO's Pandemic Influenza Severity Assessment uses as attendant properties the transmissibility, impact (e.g., effect on hospitalisations), and disease severity as input to compute severity [28], which has been successfully adapted to Covid-19 recently as well (e.g., [29]). The details are actively being investigated and therefore an attempt to formally characterise it would be part of ongoing research rather than only a representation challenge for ontologies, and thus, as temporary representation, a single property with an output value may be the least-worst option.

In sum, there seem to be nine properties, of which six are at least mandatory, if not also the necessary and sufficient conditions for something to be a pandemic. They are summarised in Table 3. There are indeed vague properties because either humans have not determined the boundaries or it is inherently difficult to measure. This will make it practically difficult to determine when something is a pandemic, especially because the fuzzy properties may generate multiple optima/minima under which a certain situation is classified as being a pandemic or not. For SARS-CoV-2 and Covid-19, for instance, at least early in 2020, it easily ticked all those boxes and so a pandemic ontology with a suitable automated reasoner would have classified the situation we were in, as a pandemic. But now, in early 2022 after the first omicron variants of concern? Of those properties, numbers 4 and 8 much less so, and number 5, on whether there will be worse novel 'variants of concern' to come, is a key question to answer. There are still out-of-season infections as well.

5. Discussion

Based on the analysis, the guiding questions posed in the Introduction can be answered at least partially. Regarding what a pandemic is and what properties it has, this has been made specific to some extent, but for as long as some of the key characteristics are imprecise due to incomplete data and knowledge by the domain experts, it is not going to be resolved with more ontological analysis. It may be possible still to capture some of it with fuzzy representation [30] and reasoning [31] if one were to be willing to use data properties in an ontology. An informal visualisation summarising pandemic, attendant entities, and vagueness is included in Fig. 1. There is a multitude of ways to formalise it, be it according to a particular modelling style [23] or ontology ecosystem like the OBO Foundry [32], and choice of logic to not only capture the crisp, declarative, static knowledge, but also the fuzzy and temporal constraints. The temporal constraints in Fig. 1 were taken from the TREND language that uses the $\mathcal{DLR}_{\mathcal{US}}$ Description Logic as foundation [33].

Concerning alignment to a FO, it appeared that the

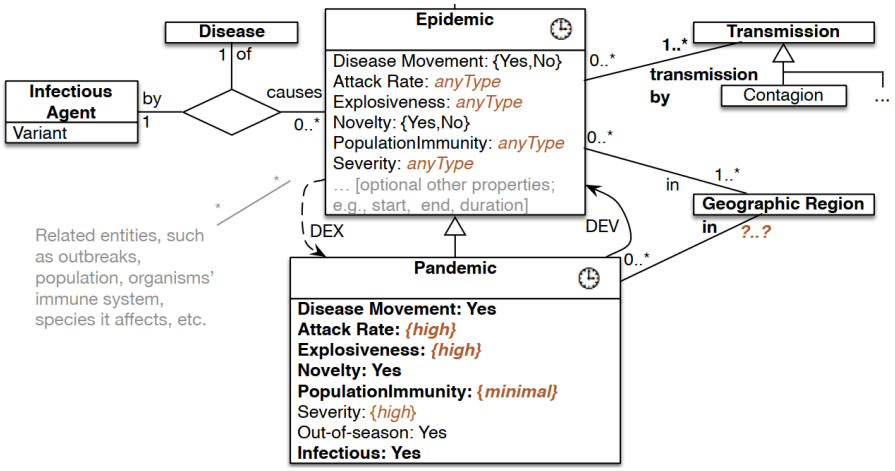

Figure 1: Informal sketch about pandemic in extended UML notation, where the knowledge is shown in very condensed format, such as depicting derived attributes that are likely to be represented formally with a set of axioms for each one (rather than as OWL data properties). Text in orange and italics: currently un- or underspecified by domain experts and their values may change over time; text in grey: content suppressed from the figure to avoid clutter; 'anyType': a suitable entity type; bold: mandatory (possibly essential) element; clock: temporal entity; DEX (+dashed): "An epidemic may also become a pandemic."; DEV (+solid): "Each pandemic must evolve to epidemic ceasing to be a pandemic.".

two with guidance were at least as good as the many hours spent trying to understand the other FOs and their documentation. The little tooling support [3, 4] reduced the time to categorise pandemic to 1-2 minutes. Assessing more FOs did not help gain more insight, other than solidifying the divide between process vs. event. In addition, the absence of guidance entails a lack of quality control mechanisms to ascertain correctness of the alignment, which does not induce confidence in the process compared to answering questions in a decision tree. This is exacerbated by the observation that for some FOs, multiple points of alignment were defensible. A 'contact the authors' guideline is undesirable as a general strategy for domain entity alignment. As also noted in [34, 5], FO tooling is needed for use and quality and, indeed, should be considered as a service to potential and actual FO users.

6. Conclusions

The assessment of domain and foundational ontologies and literature revealed that pandemic is an event (likely also an accomplishment) that unfolds in time. To be classified as a pandemic, there are a number of features from mostly attendant universals, such as the infectious agent and the disease it causes, that have to be satisfied. They are not all crisp properties and those imprecise boundaries have not all been defined. This hampers defining pandemic. Consequently, it suggests why it is difficult to determine when exactly a pandemic starts and ends.

As future work, it would be useful to develop a so-called application ontology that then can drive a simulator to model the interactions between the parameters to gain more insight and therewith improve the specification of what a pandemic is. It also would be useful if domain experts could precisiate the vague properties.

References

[1] J. Piret, G. Boivin, Pandemics throughout history, Frontiers in Microbiology 11 (2021) 631736.

[2] D. Morens, G. Folkers, A. Fauci, What is a pandemic?, J. of Infectious Diseases 200 (2009) 1018–1021.

[3] C. M. Keet, M. T. Khan, C. Ghidini, Ontology authoring with FORZA, in: Proc. of CIKM'13, ACM proceedings, 2013, pp. 569–578.

[4] C. Emeruem, C. M. Keet, Z. C. Khan, S. Wang, BFO Classifier: Aligning domain ontologies to BFO, in: FOUST-VI: 6th Workshop on Foundational Ontology, volume in print of *CEUR-WS*, 2022. 15-19 August 2022, Jönköping, Sweden.

[5] B. Smith, Ontology as product-service system: Lessons learned from GO, BFO and DOLCE, in: A. D. Diehl, et al. (Eds.), Proc. of ICBO 2019, volume 2931 of *CEUR-WS*, 2019, pp. B.1–9.

[6] H. Kelly, The classical definition of a pandemic is not elusive, Bulletin of the WHO 89 (2011) 540–541.

[7] D. Copp, What collectives are: Agency, individualism and legal theory, Dialogue 23 (1984) 249–269.

[8] S. Babcock, J. Beverley, L. Cowell, et al., The infectious disease ontology in the age of COVID-19, J. of Biomedidal Semantics 12 (2021) 13.

[9] T. Vega, J. E. Lozano, T. Meerhoff, R. Snacken, J. Mott, R. Ortiz de Lejarazu, B. Nunes, Influenza surveillance in Europe: establishing epidemic thresholds by the moving epidemic method, Influenza and Other Respiratory Viruses 7 (2013) 546–558.

[10] D. Fleming, M. Zambon, A. Bartelds, J. de Jong, The duration and magnitude of influenza epidemics: a study of surveillance data from sentinel general practices in England, Wales and the Netherlands, European J. of Epidemiology 15 (199) 467–473.

[11] P. Doshi, The elusive definition of pandemic influenza, Bulletin of the WHO 89 (2011) 532–538.

[12] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, A. Oltramari, Ontology library, WonderWeb Deliverable D18 (ver. 1.0, 31-12-2003)., 2003. Http://wonderweb.semanticweb.org.

[13] C. Beyrer, A pandemic anniversary: 40 years of HIV/AIDS, The Lancet 397 (2021) 241/243.

[14] R. Arp, B. Smith, A. D. Spear, Building Ontologies with Basic Formal Ontology, The MIT Press, USA, 2015.

[15] Z. C. Khan, C. M. Keet, Foundational ontology mediation in ROMULUS, in: A. Fred, et al. (Eds.), Knowledge Discovery, Knowledge Engineering and Knowledge Management: IC3K'13 Selected Papers, volume 454 of *CCIS*, Springer, 2015, pp. 132–152.

[16] C. Partridge, A. Mitchell, A. Cook, D. Leal, J. Sullivan, M. West, A Survey of Top-Level Ontologies - to inform the ontological choices for a Foundation Data Model, Technical Report, The Construction Innovation Hub, Centre for Digital Built Britain, 2020.

[17] S. de Cesare, C. Partridge, BORO as a foundation to enterprise ontology, Journal of Information Systems 30 (2016) 83–112.

[18] H. Herre, General Formal Ontology (GFO): A foundational ontology for conceptual modelling, in: R. Poli, M. Healy, A. Kameas (Eds.), Theory and Applications of Ontology: Computer Applications, Springer, Heidelberg, 2010, pp. 297–345.

[19] I. Niles, A. Pease, Towards a standard upper ontology, in: C. Welty, B. Smith (Eds.), Proc. of FOIS'01, IOS Press, 2001.

[20] A. B. Benevides, J. Bourguet, G. Guizzardi, R. Peñaloza, J. P. A. Almeida, Representing a reference foundational ontology of events in SROIQ, Applied Ontology 14 (2019) 293–334.

[21] G. Guizzardi, A. B. Benevides, C. M. Fonseca, D. Porello, J. P. A. Almeida, T. P. Sales, UFO: unified foundational ontology, Applied Ontology 17 (2022) 167–210.

[22] R. Mizoguchi, YAMATO: Yet Another More Advanced Top-level Ontology, in: Proceedings of the Sixth Australasian Ontology Workshop, Conferences in Research and Practice in Information, CRPIT, 2010, pp. 1–16. Sydney : ACS.

[23] P. R. Fillottrani, C. M. Keet, Dimensions affecting representation styles in ontologies, in: Proc. of KGSWC'19, volume 1029 of *CCIS*, Springer, 2019, pp. 186–200.

[24] M. E. Martinez, The calendar of epidemics: Seasonal cycles of infectious diseases, PLOS Pathogens 14 (2018) e1007327.

[25] T. Mossakowski, M. Codescu, F. Neuhaus, O. Kutz, The Road to Universal Logic–Festschrift for 50th birthday of Jean-Yves Beziau, Volume II, Studies in Universal Logic, Birkhäuser, 2015.

[26] D. Adam, A guide to R – the pandemic's misunderstood metric, Nature 583 (2020) 346–348.

[27] P. Fine, K. Eames, D. L. Heymann, "herd immunity": a rough guide, Clinical infectious diseases 52 (2011) 911–916.

[28] WHO, Pandemic influenza severity assessment (PISA), Technical Report WHO/WHE/IHM/GIP/2017.2, World Health Organisations, 2017. URL: https://apps.who.int/iris/handle/10665/259392.

[29] L. Domegan, P. Garvey, M. McEnery, R. Fiegenbaum, E. Brabazon, K. I. Quintyne, L. O'Connor, J. Cuddihy, J. O'Donnell, Establishing a COVID-19 pandemic severity assessment surveillance system in Ireland, Influenza and Other Respiratory Viruses 16 (2022) 172–177.

[30] F. Bobillo, U. Straccia, Fuzzy ontology representation using OWL 2, Int. J. of Approximate Reasoning 52 (2011) 1073–1094.

[31] F. Bobillo, M. Delgado, J. Gómez-Romero, Delorean: A reasoner for fuzzy OWL 2, Expert Systems with Applications 39 (2012) 258–272.

[32] B. Smith, M. Ashburner, C. Rosse, et al., The OBO Foundry: Coordinated evolution of ontologies to support biomedical data integration, Nature Biotechnology 25 (2007) 1251–1255.

[33] C. M. Keet, S. Berman, Determining the preferred representation of temporal constraints in conceptual models., in: Proc. of ER'17, volume 10650 of *LNCS*, Springer, 2017, pp. 437–450.

[34] C. M. Keet, Z. C. Khan, Foundational ontologies: From theory to practice and back, Journal of Knowledge Structures & Systems 3 (2022) 67–71.