Healthcare Transformation with Informatics and Artificial Intelligence
J. Mantas et al. (Eds.)

doi:10.3233/SHTI230546

What Kind of Ontologies Do We Need in the Biomedical Domain?

Marko MILETIC[a] and Murat SARIYAR[a,1]
[a] *Bern University of Appl. Sciences, Switzerland*

Abstract. We tackle the question as to what sort of ontologies we primarily need in the biomedical domain. For this purpose, we will first provide a simple categorization of ontologies and describe an important use case related to modeling and documenting events. Then, the impact of using upper-level ontologies as a basis to address our use case will be shown in order to derive an answer to our research question. Although formal ontologies can serve as a starting point to understand conceptualization in a domain and facilitate interesting inferences, it is even more important to account for the dynamic and changing nature of knowledge. Being unconstrained by pre-defined categories and relationships can facilitate timely enrichment of a conceptual scheme and provide links and dependency structures in an informal manner. Semantic enrichment can be achieved by other mechanisms such as tagging or the creation of synsets as, for example, provided in WordNet.

Keywords. Upper-level ontologies, semantic interoperability, BFO, UFO

1. Introduction

It seems to be common sense that semantic interoperability in the biomedical domain will benefit from establishing ontologies. If we are still at the level of vocabularies or taxonomies, that's just a lack of knowledge and resources. In the end, formal ontologies will facilitate many semantic and automatic reasoning challenges. This is a convincing story. A first crack in this looming promise is the ontology notion itself. There is no consensus on an authoritative definition that would guide the practice, which is not only a result of different domains using that notion. In computer science, Gruber's famous phrase "an explicit specification of a shared conceptualization" [1] leaves it open whether a taxonomy is already an ontology or if additional sorts of relations beyond the is_a-relation and axioms are necessary to form an ontology. If not stated otherwise, we will use the term "ontology" in a broad sense, covering many sorts of categorizations.

Shirky formulated an important critique on the overuse of ontologies, which Gruber discussed as well [2]. While his primary focus was on annotations and categorizations on the internet, his insights are applicable to other settings as well. For instance, he described how user-generated tags and links are more effective than pre-determined hierarchical categories when it comes to guiding general searches for information. As information needs are highly context-dependent, it is not useful to categorize, for example, books under entertainment rather than under science or culture. In addition to that, the dynamic nature of the internet poses a challenge for maintaining a constant

[1] Corresponding Author: Murat Sariyar, Bern University of Applied Sciences, Quellgasse 21, CH2502 Biel/Bienne, Switzerland; E-mail: murat.sariyar@bfh.ch.

information dependency structure. One consequence is the requirement to provide criteria as to when ontologies might be useful, for instance: a small corpus, formalizable categories, stable and restricted entities, and clear edges (relations). Another consequence is to forego formal hierarchical methods and instead opt for flexible classification schemes, like WordNet [3].

Here, we tackle the question as to what sort of ontologies we primarily need in the biomedical domain. For this purpose, we will first provide a simple categorization of ontologies and describe an important use case related to modeling and documenting events. Then, the impact of using (different) upper-level ontologies as a basis to address our use case will be shown in order to derive an answer to our research question. Such an overall perspective is still rare in ontology development. We will conclude by discussing the general implications of our results to ontologies in the biomedical domain.

2. Methods

Four levels of ontologies can be discerned in computer science and applied ontology: (i) upper-level ontologies such as the Basic Formal Ontology (BFO [4]) and the Unified Foundational Ontology (UFO [5]). (ii) General domain-level ontologies for a research field. (iii) Specific domain-level ontologies for subject matters such as sensory data in a home environment. (iv) Lightweight ontologies for subject matters that are not as formalized as the former ones. The fourth category covers many different kinds of modeling approaches, for example, classification, thesauri, and taxonomies, which do not represent a description of a domain using classes, different relationships, and logical axioms. In most cases, realism is assumed in all categories of ontologies, which means that the concepts reflect reality and not a construction thereof. When information is to be modelled, this might cause a problem, as understanding of concepts is not just a context-dependent but also a mind-dependent process (information is the result of interpretation).

As a use case for adapting an ontological approach, documenting events in interoperative monitoring during neurosurgery will be used, for which we previously developed a formalized ontology [6]. An event is related, for instance, to the amplitude or latency change in signal measurement during surgery. We developed our own ontology instead of adapting an existing one for several reasons. The main reason was that the most promising alternative, the Ontology of Adverse Events (OAE [7]), had many conceptual issues that became apparent after reading its definition of adverse events as “a pathological bodily process that occurs after a medical intervention”. Even tough, equating events with processes might be suggested by the BFO founders [8], there is a need to differentiate between basic and complex events due to questions of identity criteria and whether events are particulars or universals, which was crucial for our use case of documenting events, as it mostly requires exact reference to events as particulars. These sorts of questions are related to fact, that there are no guidelines on how to describe events, which triggered a fundamental concern about the appropriateness of using ontologies at all in our setting. We have assumed that enriching the documentation of events with content-related relationships such as causally-related-to or occurs-in would be useful. To address this concern, we have posed the question of which level of ontology development (ii)-(iv) should be targeted, with the final category serving as a placeholder for no formal ontology. For this decision, we investigated top-level ontologies on level (i) in their stability and relatedness. Besides BFO und UFO, we also considered the ontology of EJ Lowe as a pure philosophical outcome guiding UFO [9].

3. Results

We made two important observations regarding all three upper-level ontologies: (a) They have fundamentally different distinctions in their categories despite the fact that all assume the perspective of realism and descriptive metaphysics; (b) They had different number of changes throughout their history, but always to such an extent, that the potential impact for the practice was significant. Starting with the details of the first point: In BFO the central distinctions is between continuants and occurrents within one categorization; UFO has three separate categorizations for endurants (UFO-A), perdurants (UFO-B), and intentional as well as social entities (UFO-C); Lowe uses the distinction between universals and individuals at the highest level. These differences associated with the fact, that BFO just includes is_a-relations, where the other ontologies also have instance_of-relations, which are delegated to domain-specific ontologies and the creation of a knowledge base by BFO. Regarding an overall ontological commitment to events, these differences suggest that incongruent perspectives need to be considered.. In BFO, events are occurrents that can be spatio-temporal regions as well as processes. In UFO, an event is suggested to be a couple <r, f> where r is a spatio-temporal region, and f is the event's focus, consisting of a collection of individual qualities that are cognitively sorted out from a scene, which is the sum of relationships that are effective for producing the events in questions. Lowe's ontology has less ontological commitment and allows to model events as (universal) properties or as concrete non-objects (tropes or modes) with no explicit spatio-temporal regions.

With respect to the changes of the upper-level ontologies, BFO was released in 4 major versions (1.0, 1.1, 2.0, and version 2020 as an ISO standard). The categorial core of BFO, resting on the distinction between continuants and occurrents, and between dependent and independent entities, has remained constant. However, in 1.1 the category of generically dependent continuant was added, allowing to represent information artefacts. For modelling events, this meant to be able to focus on documenting these events, which was the choice in our use case because thereby we could in the first instance circumvent problematic issues such as the identity and essence of events. In the UFO ecosystem many changes were developed with two major versions being 0.1, 0.2 and 1.0. As with BFO, the core remained stable, but in 0.2 categories of datatypes and processes were included, and in 1.0 the detail level and consistency were increased, e.g., by including qualia as sub-categories of abstract individuals. For events, it is important, that they can be now framed by time-intervals and have substances as participants. Lowe made also substantial changes to his ontology without giving up the core distinction between universals and individuals. For example, he switched from the primary division of particulars into objects and tropes to a division into abstracta and concreta, and under the latter objects (having criteria of identity) as well as non-objects (especially tropes) are subsumed. In such an ontology, events as tropes are better distinguished from abstract particulars such as qualia or sets. However, a determination of events as tropes with no criteria of identity (incomplete or unsaturated entities) is debatable as many philosophers such as Quine, Lewis, Kim, and Bennett oppose this view (see [10]).

These insights indicate that the conceptual hierarchy in upper-level ontologies involve some instability, at least below fundamental distinctions. One resulting problem is the challenge of mapping ontologies based on BFO 1.1 to those based on BFO 2.0, where changes such as separating processes and process boundaries and placing them at the same level as spatio-temporal regions, instead of one level below spatio-temporal regions under the concept of "processual entity," can occur. We do not need to mention the much more severe problems of mapping ontologies based on BFO to ontologies based in UFO, regardless of the versions considered. Our additional issue of having no clear guidance from upper-level ontologies for modeling events was recently acknowledged in the UFO universe. Guarino et al. [10] go a step back from formal ontologies to central questions in the form “What are events?” and “What is the referential mechanism that is in play when we describe an event?”, making clear how important it is to consider the whole context of an event. Based on the forgoing, we conclude that the instability and the lack of contextual considerations in upper-level ontologies as well as in domain-specific ones at level (ii) and (iii) based on them make them not only questionable for our use case, but for most application in the biomedical domain. There is a need for lightweight and adaptable approaches to knowledge organization at the level (iv) of our categorization, which also avoids time-consuming and costly process of creating and maintaining a formal ontology. Semantic enrichment can be achieved by other mechanisms such as tagging or the creation of synsets.

4. Discussion

Here, we used observations on three important upper-level ontologies to justify a relativization of formal ontologies in the biomedical domain. Even if formal ontologies could be a starting point to understand the conceptualization in a domain and to allow for interesting inferences, it seems much more important to account for the dynamic and changing nature of knowledge. Being unconstrained by pre-defined categories and relationships is helpful in timely enriching a conceptual scheme and providing links and dependency structure in an informal way. Context-dependency can hardly be captured by a formal structure, which even the authors of UFO realized in the case of events by providing a substantial account of philosophy of events. We are aware of the fact, that these considerations require more room to be developed fully. These two topics will be extended in the near future: the impact of event theories on ontology development and the need to include a constructivist-representational element in (lightweight) ontologies.

References

[1] Gruber TR. A translation approach to portable ontology specifications. Knowledge acquisition. 1993 Jun 1;5(2):199-220.

[2] Gruber T. Ontology of folksonomy: A mash-up of apples and oranges. International Journal on Semantic Web and Information Systems (IJSWIS). 2007 Jan 1;3(1):1-1.

[3] Miller GA. WordNet: a lexical database for English. Communications of the ACM. 1995 Nov 1;38(11):39-41.

[4] Arp R, Smith B, Spear AD. Building ontologies with basic formal ontology. Mit Press; 2015 Jul 31.

[5] Guizzardi G, Botti Benevides A, Fonseca CM, Porello D, Almeida JP, Prince Sales T. UFO: Unified foundational ontology. Applied Ontology. 2022 Jan 1;17(1):167-210.

[6] Neuenschwander S, et al. Developing an Ontology for Documenting Adverse Events While Avoiding Pitfalls. Informatics and Technology in Clinical Care and Public Health. 2022 Jan 14;289:166.

[7] He Y, Sarntivijai S, Lin Y, Xiang Z, Guo A, Zhang S, Jagannathan D, Toldo L, Tao C, Smith B. OAE: the ontology of adverse events. Journal of biomedical semantics. 2014 Dec;5(1):1-3.

[8] Smith B. Classifying processes: an essay in applied ontology. Classifying Reality. 2013 Apr 17:101-26.

[9] Lowe EJ. More kinds of being: A further study of individuation, identity, and the logic of sortal terms. John Wiley & Sons; 2015 Feb 23.

[10] Guarino N, et al. Events, their names, and their synchronic structure. Applied Ontology. 2022 Jan 1(Preprint):1-35.