Ontologies for the digital humanities: Challenges and perspectives

Emilio M. Sanfilippo
ISTC-CNR Laboratory for Applied Ontology

There is a growing interest in the Digital Humanities (DH) for the use of ontologies for data management tasks such as data access, publication, consistency checking, and integration, to mention some common examples [12, 13]. In these contexts, at the current state of the art, Semantic Web (SW) languages and technologies [8] are the resources mostly used opening new ways and challenges for handling data. Neither ontologies, nor SW resources are however “wizard sticks”, and efforts are required to meet both domain experts’ needs and applications requirements, as well as to make ontologies and tools robust and reusable items.

The purpose of this talk is to address some of the challenges emerging from the use of ontologies in the DH, focussing in particular on the conceptualization of experts’ knowledge in ontological terms. We aim in this way at individuating some peculiar conceptual dimensions that ontologies for the DH need to address or, otherwise said, at individuating some research questions for further research in the field. The analysis is partial and general at the same time in that we will leave aside, e.g., research aspects relative to the usability of ontologies (the scalability of RDF triplestores, etc.) as well as modeling challenges relative to specific domains (e.g., musicology, history of arts, philology, etc.). The hope is that the analysis can still help in identifying some interesting research questions deserving attention from the broad DH landscape interested in knowledge representation and data management.

Conceptual drift. Considering concepts from a diachronic perspective, the way in which we use them today does not necessarily correspond to the manner in which they were used

1The reader can also refer to the special issue about culture heritage on the Semantic Web journal: http://semantic-web-journal.net/blog/special-issue-semantic-web-cultural-heritage
in other epochs because of changes (drifts) in their meaning. One of the most common examples is the concept of democracy, which for ancient Greeks referred to a political form closer to what we would call nowadays participative democracy; or the concept of blanc-manger which in Middle Age culinary receipts referred to salty dishes prepared with either fish or white meat, whereas it is used today for a white pudding (a type of dessert) [1, 9].

Ontologies may need therefore to take into account the history of concepts while possibly accommodating the shifts in their meanings, especially when data coming from different periods are to be managed (see [5, 11] for examples). This poses fundamental questions about how to conceive concepts, e.g., relative to their identity and persistence conditions. Consider the concept of blanc-manger in its double characterization as a type of salty dish and as a type of dessert. How many concepts are there? If the answer is one, this means that we allow concepts to change while remaining essentially the same, similarly to the way in which an object can change while remaining identical to itself. On the other hand, if the answer is two, it means that we have some concerns about concept change. There is no easy solution and it is clear that, to face these sorts of challenges, we first need a theory of concepts making explicit how they are characterized and how they can persist in time.

Finally, recall that ontologies do not force the adoption of one concept at the expenses of others. For instance, should we distinguish between different concepts for blanc-manger, we do not necessarily need to pick up the “correct” one while discarding the other. Pluralistic approaches allow indeed characterizing the differences among multiple resembling concepts capturing at the same time their relations (see [7] for an example about the notion of artefact). Research efforts in this direction are needed.

**Uncertain information.** Data in the DH can be partial, uncertain or even unknown, especially about past events. Examples are dates (e.g., when someone was born/died), the geolocation of places, the authorship of literary or musical works, to mention only few cases. In addition, domain experts may express their claims in probabilistic terms, e.g., when saying that it is only probable that a person is the author of a document [4].

How to make sense of these phenomena in ontologies is open for research. Recall that in knowledge representation, epistemic logics are used to make explicit what agents believe or know [14], and fuzzy logics are used to ascribe probability values to logical propositions [3]. One may therefore exploit these sorts of formalisms to represent scholars’ hypotheses about the investigated phenomena and their degrees of probability (see, e.g., [10]). Existing research also explores the use of these approaches in computational logics (e.g., [2]). At the same time, however, standard SW formalisms and state of the art tools rely on plain classical logic [8], hence they do not allow representing probability and lack the (logical) operators for agents’ epistemological status. Efforts are needed to exploit formal languages other than those standardly adopted in SW frameworks.

Research works going in this direction while adopting classical logic exist although they are still preliminary (e.g., [4, 15]). For instance, one should investigate how to conceive beliefs independently from the mental states of specific agents, what are their persistence
conditions, and how multiple beliefs about the same phenomena can co-exist, even if they are incompatible with each other (e.g., when they ascribe incompatible birth dates to the same person). From this perspective, the modeling of probability values is likely to be left aside for technical reasons; this could be seen as a trade-off between the expressivity of the chosen modeling languages and their use for practical application goals.

**Interpretations and intentional properties.** It is common for scholars in DH projects to characterize domain entities via *intentional properties*, roughly speaking, properties that exist only because someone has ascribed them to the entities at stake. For instance, in musicological contexts, it is an open question what is the genre of a musical composition, and disagreements can arise about whether a piece is, say, a mazurka or something else.

Differently from what was discussed in the previous paragraph, the attribution of intentional properties, in this case, does not depend on the lack of data. It is rather due to the fact that scholars elaborate *interpretations* about domain entities resulting in such attributions. An example is the interpretation of a painting or a literary text: a scholar may interpret a character appearing in a novel from a psychoanalytic perspective attributing it properties which may not be even taken into account by a scholar analysing the same character but from a Marxist perspective. It should be however clear that there are strong relations with what discussed above in that an interpretation may be conceived as a type of belief, a further motivation to take beliefs seriously into account for ontology design.

Concerning intentional properties, what is interesting from an ontological perspective is that they bear a *relational nature* since they depend on both their bearers (e.g., a musical composition) and the scholar(s) who has made the claim about them. For instance, one cannot simply say that a composition has a certain genre but that it has that genre *because of, according to* a scholarly belief/agent.

Research efforts in these directions are required. Apart from the analysis and representation of beliefs (see previous paragraph), it is necessary to investigate the nature and persistence conditions of intentional and relational properties (see [6]). Intuitively, one may wish to say that a composition has a genre according to a belief even when the agent who has formulated the belief does not exist anymore, a consideration opening the way to investigate the relation between a belief, its creator, and the property it attributes to the entity the belief is about.

**References**


---

See the entry *Genre* by Jim Samson on the Grove Music Online.


