Formal Ontology in Information Systems
S. Borgo et al. (Eds.)

doi:10.3233/978-1-61499-910-2-1

Some Open Issues After Twenty Years of Formal Ontology

Stefano BORGO [a,1] and Pascal HITZLER [b]
[a] *Laboratory for Applied Ontology, ISTC-CNR, Trento, Italy*
[b] *Data Semantics (DaSe) Laboratory, Wright State University, Dayton, OH, USA*

1. Introduction

Twenty years ago, the introduction of the first conference on Formal Ontology in Information Systems, FOIS 1998, marked the state of the art and the problems that research on ontology needed to face to have an impact on applications [9]. At that time, computational ontologies were emerging as a young research topic whose aim, broadly understood, was to overcome the traditional problems in information understanding, management and sharing. That conference was the first one addressing this topic under an interdisciplinary perspective, explicitly acknowledging the role of ontology as a branch of philosophy, as well as linguistics and cognitive science. Later on, while international events and publication venues were opening their initiatives to research in applied ontology, researchers were struggling with many open questions: from the understanding of the role of ontological systems to the search of methodologies for ontology construction, from the distinction of ontology typologies to the identification of an ideal (logical) language and formalization level.

In the FOIS 1998 introduction, Guarino provided a definition of formal ontology that influenced and shaped the domain thereafter. Building on Gruber's intuition of conceptualization, Guarino proposed to see an ontology as a "logical theory accounting for the intended meaning of a formal vocabulary." This view, further elaborated in later works [11], and still debated today, is fairly well accepted in the area of formal ontology but not in the ontology community at large where a less constrained view has been practically endorsed. Today it is common to call ontology any logical theory that includes a taxonomy and that is written in a computational (and often decidable) language like the Web Ontology Language (OWL) [16].

A second topic raised in [9] is the ideal level of formalization of an ontology. Since a formal ontology is a logical theory about the world (more precisely, a conceptualization of a part of the world) and logic is neutral with respect to the nature and structure of the world, every piece of information about the nature and the structure of the world has to be explicitly coded into a set of logical axioms. This raises two problems. On the one hand, the number of sentences to be added in the logical system turns out to be quite

[1]Corresponding Author: Stefano Borgo, LOA ISTC-CNR, Trento, Italy; E-mail: stefano.borgo@cnr.it

large and depends on the granularity with which one looks at the world. On the other hand, the result depends on the type of language one uses. If we use a decidable logical language, for technical reasons the capacity of the language to express the needed information is limited and the logical theory will not be able to include basic information (for instance, these languages have limitations in coding ternary relations like "z is between x and y" or "x counts as y at time t"). Instead, if we use expressive logical languages, which are undecidable for technical reasons, the logical theory becomes soon unmanageable, consistency cannot be ensured with standard techniques[2], and reasoning with it is difficult.

Another issue regards the proposal presented in [9] to separate ontologies in four types: top-level, domain, task and application ontologies. Each type collects ontologies that aim to model different aspects of the world conceptualization. Top-level ontologies are devoted to "describe very general concepts like space, time, matter, object, event, action, etc., which are independent of a particular problem or domain"; domain ontologies "describe [...] the vocabulary related to a generic domain (like medicine, or automobiles)"; task ontologies "a generic task or activity (like diagnosing or selling)"; finally application ontologies "describe concepts depending both on a particular domain and task, which are often specializations of both the related ontologies. These concepts often correspond to roles played by domain entities while performing a certain activity, like replaceable unit or spare component." (citations from [9]). The theoretical distinction had only a limited impact since today domain ontologies and application ontologies are largely used as synonyms, task ontologies are rarely addressed and the most recognized (and perhaps useful) separation is between top-level ontologies and application ontologies. An important factor leading to this result is that the practitioner often finds it hard to distinguish the classification of entities by type and by role. After all, when it comes to application concerns, the ontology is naturally influenced by the observer's perspective and this freezes the entities into their contextual roles. Once an entity is seen as maintaining the same role across any scenario in the application at stake, the importance of the theoretical distinction between the entity and its role fades away.

Twenty years on, many things have changed. Doubts about the possibility of Artificial Intelligence (AI) have disappeared, the Semantic Web (SW) has become a reality, perhaps in a less ambitious form where ontologies are "knowledge graph schemas", HTML changed the traditional meaning of link, and open (linked) data that of privacy. In this same period, formal ontology enhanced both AI and SW, and definitively changed the meaning of sharing, which evolved from a distribution of mere data, to a distribution of data linked to independent resources (from taxonomies with implicit semantics, like schema.org, to dedicated ontologies). In this temporal frame several steps ahead have been made in formal ontology. Today this is a well known research area and has set important results to clarify the relevance, the interaction across and the formalization of several core theories like those of space and time, identity, essential properties and qualities, roles, parthood, inheritance, dependence and constitution as witnessed by papers published in the FOIS conferences and Applied Ontology journal.[3] It also has practi-

[2]Only the first-order theory of the DOLCE ontology [21] has been proven to be consistent via the use of a sophisticated logical technique [19].

[3]http://content.iospress.com/journals/applied-ontology

cal applications in several domains, see for instance [18,23,6,8] – just to point to a few papers in the last FOIS conferences.

The application of formal ontology has been sometimes successful and at times less satisfactory for a number of reasons. Among the major problems is the lack of a clear list of competences that are required to do ontological analysis and to build ontologies [25]. Unfortunately, today one cannot establish whether a practitioner has the needed expertise to work in formal ontology by looking at his/her education or training history as one does for, say, mathematics or medicine. Another important bottleneck for the establishment of a satisfactory level in ontology construction is the high variety of systems that are called ontologies in the different communities. Today there are clear parameters to evaluate databases or logical languages but ontology research has not yet identified a comparable set of criteria for assessing the result of an ontological analysis or to evaluate an ontological system [26].

2. Some Open Problems in Formal Ontologies

The history and discussion outlined above naturally gives rise to research directions which appear to be central to the future development of the field, and in particular its ability to stay current with ongoing developments in neighboring fields and applications. We first look at some theoretical issues and then at practical ones. Our goal is to stimulate the ontology community to explicitly investigate these aspects.

2.1. Theoretical Issues

One leitmotif of ontology research in these twenty years has been and still is the need to explicitly state principles and guidelines for ontology construction and exploitation. Guidelines for the use of a formal ontology are important since they can ensure consistency and the correct adaptation of the ontology to the application domains. For instance, each formal ontology allows the user to extend the system to include new domain categories. Since this means to choose what to introduce and how to do it, the process is prone to errors, not much from the logical viewpoint but from the ontological one. Guidelines are helpful to ensure that the users' extensions are compliant with the general structure and perspective of the initial ontology. Furthermore, existing information on how to ontologically analyze an entity and a list of actual examples of analyzed and classified entities and relations, help to understand the formal ontology and to use it homogeneously, especially in the delicate phase of ontology population, that is, when the individuals are classified into the system.

The role of philosophically inspired principles is even more critical in this research area since their consistent and integrated use is what distinguishes formal ontologies from traditional classification systems. For instance, while a top-level ontology is a classification system that deals with general domain-independent categories only, a foundational ontology is a top-level (formal) ontology that has been built and motivated by the upfront and explicit choice of its core principles. Principles are about fundamental choices: the understanding of space/time, the relationship between entities and space/time, the existence of objects and/or events, the existence of abstract entities, of possibilia, of types of properties, the relationship between objects/events and their properties, the identity conditions, the dependence relationships and so on.

Some of the existing formal ontologies go further and explicitly adhere to specific philosophical schools. For instance, a formal ontology like BFO [1] is an attempt to translate an interpretation of scientific realism into logical form, while the formal ontology GFO [15] explicitly relies in a form of conceptual realism called integrative realism. We note that today no philosophical school recognizes these systems as truly representative of (a form of) realism.[4] Other ontologies are only indirectly inspired by some philosophical school, and prefer to make more pragmatic choices. This is, for instance, the case of YAMATO [22] which, while vaguely realist in spirit, avoids any explicit commitment to that or other philosophical views.

Ontologies that roughly fall within a conceptualist approach exist as well, for instance UFO [14], an attempt to unify DOLCE [21] (discussed below), GFO, and OntoClean [12]. Also the ontology GUM [2] (the outcome of efforts motivated by linguistic analysis) is better understood within this perspective. These formal ontologies do not explicitly commit to a specific philosophical school and their orientation towards conceptualism (and perhaps conventionalism) is a consequence of the initial motivations for their construction more than the result of an upfront philosophical choice. This observation characterizes most of the remaining ontological systems as they tend to recognize different ontological viewpoints. GUM is open to the integration of multiple modeling perspectives (a natural outcome of the focus on linguistic semantics). Similarly, DOLCE [21] is presented as an ontological system within a library of co-existing formal ontologies. DOLCE falls into this class also for another reason. The ontology, which was initially inspired by linguistic and cognitive arguments, has been revised to foster interoperability by adopting logical formulas that are equally acceptable by users with different philosophical commitments [5]. Indeed, it is clear that BFO and DOLCE, to name two systems that make similar philosophical choices on several aspects, interpret the role of formal ontology in information science in radically different ways. BFO requires all its users to embrace the realist philosophical view in order to ensure interoperability. DOLCE requires only ontological coherence and lets users maintain their world view by showing how to integrate different modeling perspectives within a single formal ontology.

Finally, there are other systems that have been developed as top-level ontologies and subsequent efforts tried to make them closer to principled formal ontologies. This is the case of SUMO [27] which was developed by assembling off-the-shelf formal theories that together could cover the most general categories. An analysis of the philosophical stand of SUMO was taken only ex post (without a clearly established result since the system seems to follow slightly different views in different parts of the ontology). Another example is given by the ISO 15926 standard [3] which, developed from practical considerations, relies on a so-called 4D top-level ontology.[5] This is an interesting but today less explored ontological position.

[4]It is also unclear whether these philosophical stands have a technical impact. BFO and GFO's logical theories are quite weak on this aspect. See [7] for a technical discussion of this topic.

[5]Roughly, it admits the existence of events (things that necessarily exist and develop in time) but not of objects as traditionally understood.

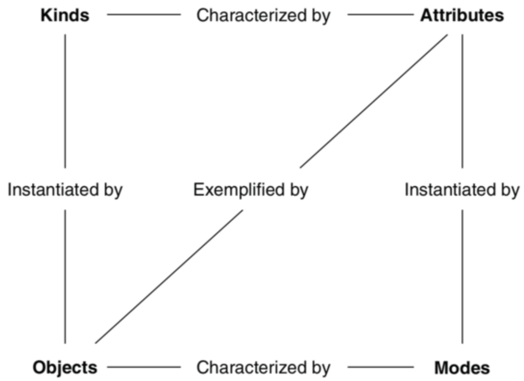

Figure 1. The four-category ontology [20].

Given this brief outlook of some formal ontologies, one wonders whether their different stands have a real impact in the use of the formalized systems. Here things are less clear. Generally speaking, where something cannot be modeled directly because of an ontological distinction (e.g. the sharp distinction between material and information objects blocks the simple representation of common-sense objects like a book), these ontologies propose reformulation patterns that are often successful (e.g. the distinction of the book as the material entity and the book as the information entity among which a representation relation holds). Yet, this method cannot always work due to the idiosyncrasy of some application domains and the complexity of some concepts used in applications. To state it briefly, all these ontologies explicitly state their core principles but then allow the user to add arbitrary categories even though these are not ontologically justified or justifiable. While this is built-in in ontologies like DOLCE and GUM, unprincipled extensions should be alien to formal ontologies that adhere to some philosophical school. For the latter systems, the actual exploitations of the ontology in application domains are philosophically wrong but practically accepted, and since an applied ontology has its *raison d'être* in its use, they make a virtue out of necessity.

Looking at the different formal ontologies as logical systems, one quickly concludes that they are all pairwise inconsistent. Yet, if we look at the philosophical principles that inspire the ontology organization and construction, we find that many core assumptions are shared as most of these formal ontologies embrace the neo-Aristotelian ontology structure (Fig. 1). The differences arise on how to understand even the basic elements in this structure: the identity criteria of objects, the existence of abstracts, the ontological nature of events, how to understand properties.

Are these differences relevant to the construction of ontological systems whose goal is to overcome interoperability problems? Unfortunately, we lack scientific studies to definitely answer this question. For this reason, we hope the ontology community will pay more attention to this problem. Indeed, the interplay between philosophical principles and standard application concerns is still largely unexplored.

In these twenty years, formal ontologists looked at philosophy to get valuable principles for ontological analysis and construction. The effort has been on the development of ontological systems that guarantee coverage while being conceptually clear and logically consistent. Experience has shown that this is a complicated effort and only a few systems have emerged. With the introduction of specific languages and software tools (e.g. DOL [24]), it will become easier to design and build good ontological systems. The availability of dedicated software tools will have three positive impacts: (a) the ontologist will be able to test different combinations of principles (via their logical formalization) to verify the most suitable systems for ontology clarity and interoperability concerns; (b) the need to follow philosophical stands will reduce since the choice of principles and the verification of their logical consistency can be done on different bases, and the concerns will finally concentrate on clarity and interoperability; (c) ontological systems will be able to integrate results from linguistics and cognitive science to test and enhance transparency and communication capacities. The consequence is that it will finally become possible to define benchmarks in formal ontology which are based on verifiable ontological consistency, conceptual transparency and interoperability capacities. At the same time philosophical principles, freed from the constraints and divisions they bring with them today, will continue to have their important role in formal ontology.

Will this be enough to identify applied ontology as a scientific discipline that stretches across philosophy and application domains while remaining independent from them? Twenty years later, it appears the time is ripe to verify this hypothesis and to start a new phase in formal ontology.

2.2. Practical Issues

From the practical viewpoint, there is no doubt of the value of ontology modeling based on philosophical principles. However, it is clear that this approach raises several problems. One of them is the entrance barrier for newcomers: formal ontologies based on principles are hard to understand in particular because some of the concepts and foundations, and their implications, appear to be difficult to grasp without a substantial background in their underpinnings. Another related problem is the high cost in terms of time and expertise that managing and reusing such modeling systems require. These ontologies are prohibitive to adapt (e.g. integrating specialized modules, pruning irrelevant parts, adding new stakeholders' views) in many contexts where there are time constraints and specialized personnel is not available. Pragmatic requirements therefore have been calling for trade-offs between the foundational and the ad-hoc approaches to ontology modeling. Foundational research is expected to guide developments on this trade-off but little progress has been done in these years, a notable but limited exception being the development of ontology design patterns [17]).

While there is agreement that highly formal ontology modeling leads more likely to robust ontologies, which by construction are reusable in many contexts, their generation also requires significant efforts. On the other hand, ad-hoc computational ontologies are rather cheap to produce, but tend not to follow quality principles and are extremely limited in terms of reusability,[6] which in turn means that it often seems to be easier to make a new ontology rather than attempt to reuse or modify such an ad-hoc ontology. One important reason is that ad-hoc ontologies tend to be developed taking a particular viewpoint or purpose, and this makes it hard to adapt them to different contexts or uses. Consequently, ad-hoc ontologies have a significant cost factor in terms of time and ex-

[6]The common observation that many if not most ontologies are not really being reused in practice, may be attributed to their ad-hoc nature. This means that the underlying ontological commitments were often not informed by foundational principles.

pertise required when they are to be updated, modified, or repurposed. These observations – where both extremes come with high cost – suggest that there should be a soft spot, a trade-off point where the cost-benefit ratio is optimal. Of course, it might be that the optimal trade-off point depends on the use and reuse context, i.e. intended current and future use, of the ontology. We still do not have an active line of research on how to achieve such a favorable trade-off, nor even a list of quality metrics to evaluate such result.

A trade-off should also be sought on the continuum between a strong and comprehensive versus a weak and shallow logical axiomatization of the ontology. This relates to both the choice of logical knowledge representation language used for the encoding (which sometimes is driven by application constraints), and to the question of how deeply axiomatized should the notions be. Roughly, this means how many axioms should be added to the system and how fine-grained and detailed an axiomatization should be. In terms of language, both description logics and rules have been posited as favorable for several reasons [28], however hardly any research has been done on this question. In terms of depth of axiomatization, there again appears to be a trade-off, as formal semantics is made to restrict meaning, i.e., more axioms lead to a narrower scope for the definitions as these are constrained by the axiomatization. Few axioms thus lead to ambiguities, while many axioms put hard constraints on reusability as they may impose too-narrow ontological commitments. As before, it seems reasonable to conjecture that there is a soft spot, a favorable trade-off between both extremes, but little research has been done on this issue.

Investigating these trade-offs is by no means a trivial or straightforward task. In particular, it requires to new dedicated approaches and extensive user studies for the development of quality metrics which remain independent of specific philosophical schools of thought.

Finally, modeling based on the formal ontology approach has to be made practicable by lowering the adoption barrier while maintaining reasonably high quality standards. This means the development of tools and modeling methodologies that ease the burden of philosophical and logical foundations. Different directions can be explored. We already mentioned the relevance for ontology research of software like DOL [24], which provides a unified metalanguage for handling the diversities across formal languages, models and specifications. One interesting advantage of this approach is the possibility to verify consistency across ontology modules facilitating reuse and adaptation. On the other hand, and perhaps more interestingly from a practical viewpoint, one can develop formal languages whose constructs are ontologically non-neutral [10]. The goal here is to have ontological distinctions built directly into the representation language so that the user does not need to code the ontological assumptions of the constructs, an ontologically subtle and error prone activity. Within this view, the approach taken by OntoUML[7] [13] is a promising line of research today. From a more theoretical viewpoint, a variety of ontologically specialized formal operators should be exploited, see e.g. [4].

Rigorous empirical research could shed light onto the effectiveness of specific methodologies and tools, and contributions to these methodologies from researchers in foundational ontology is strongly needed.

[7] http://www.menthor.net/ontouml.html

References

[1] Robert Arp, Barry Smith, and Andrew Spear. Building ontologies with Basic Formal Ontology. MIT Press, 2015.

[2] John Bateman, Bernardo Magnini, and Giovanni Fabris. The generalized upper model knowledge base: Organization and use. Towards very large knowledge bases, pages 60–72. IOS Press, 1995.

[3] Rafael Batres, Matthew West, David Leal, David Price, Katsube Masaki, Yukiyasu Shimada, Tetsuo Fuchino, and Yuji Naka. An upper ontology based on ISO 15926. *Computers & Chemical Engineering*, 31(5–6):519 – 534, 2007.

[4] Stefano Borgo, Daniele Porello, and Nicolas Troquard. Logical operators for ontological modeling. In Oliver Kutz and Pawel Garbacz, editors, *FOIS 2014 - Formal Ontology in Information Systems*, volume 267 of *FAIA*, pages 23–36. IOS Press, 2014.

[5] Stefano Borgo and Claudio Masolo. Foundational Choices in DOLCE. In S. Staab and R. Studer, editors, *Handbook on Ontologies*, pages 361–381. Springer Verlag, 2nd edition, 2009.

[6] Duarte Bruno Borlini, Souza Vítor E. Silva de Castro Leal, André Luiz and de Almeida Falbo, Ricardo and Guizzardi, Giancarlo and Guizzardi, Renata SS, Towards an Ontology of Requirements at Runtime. In *Proceedings of the Ninth International Conference on Formal Ontology in Information Systems*, pages 255–268. IOS Press, 2016.

[7] Nino B. Cocchiarella. Philosophical perspectives on formal theories of predication. Handbook of Philosophical Logic, Vol. 4, pages 254–326. D. Reidel Publishing Company, 1989.

[8] Antony Galton, Gabriel Landini, David Randell, and Shereen Fouad. Ontological Levels in Histological Imaging. In *Proceedings of the Ninth International Conference on Formal Ontology in Information Systems*, pages 271–284. IOS Press, 2016.

[9] Nicola Guarino. Formal ontology in information systems. In N. Guarino, editor, *Proceedings of the Second International Conference on Formal Ontology in Information Systems*, pages 3–15. IOS Press, 1998.

[10] Nicola Guarino. The ontological level: Revisiting 30 years of knowledge representation. In *Conceptual modeling: Foundations and applications*, pages 52–67. Springer, 2009.

[11] Nicola Guarino, Daniel Oberle, and Steffen Staab. What is an ontology? In *Handbook on ontologies*, pages 1–17. Springer, 2009.

[12] Nicola Guarino and Chris Welty. An overview on OntoClean. In S. Staab and R. Studer, editors, *Handbook on Ontologies*, pages 201–220. Springer Verlag, 2nd edition, 2009.

[13] Giancarlo Guizzardi. *Ontological Foundations for Structural Conceptual Models*, volume 05-74 of *Telematica Instituut Fundamental Reserach Series*. CTIT: Centre for Telematics and Information Technology, Enschede (The Netherlands), 2005.

[14] Giancarlo Guizzardi, Gerd Wagner, João Paulo Andrade Almeida, and Renata SS Guizzardi. Towards ontological foundations for conceptual modeling: the Unified Foundational Ontology (UFO) story. Applied ontology 10(3-4), pages 259–271. IOS Press, 2015.

[15] Heinrich Herre. General Formal Ontology (GFO): A foundational ontology for conceptual modelling. In Theory and applications of ontology: computer applications, pages 297–345. Springer, 2010.

[16] P. Hitzler, M. Krötzsch, B. Parsia, P. Patel-Schneider, S. Rudolph (eds.). OWL 2 Web Ontology Language: Primer (Second Edition). W3C Recommendation (11 December 2012), available at http://www.w3.org/TR/owl2-primer/

[17] P. Hitzler, A. Gangemi, K. Janowicz, A. Krisnadhi, V. Presutti (eds.). Ontology Engineering with Ontology Design Patterns: Foundations and Applications. Studies on the Semantic Web Vol. 25, IOS Press / AKA Verlag, Amsterdam, 2016.

[18] C. Maria Keet and Langa Khumalo. On the ontology of part-whole relations in Zulu language and culture. In this volume.

[19] Oliver Kutz and Till Mossakowski. A modular consistency proof for DOLCE. In *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*, pages 1–8, August 2011.

[20] E. Jonathan Lowe. *The Four-Category Ontology. A Metaphysical Foundation for Natural Science*. Oxford University Press, Oxford, 2006.

[21] Claudio Masolo, Stefano Borgo, Aldo Gangemi, Nicola Guarino, Alessandro Oltramari, and Luc Schneider. DOLCE: a descriptive ontology for linguistic and cognitive engineering. *WonderWeb Project, Deliverable D17 v2.1*, pages 1–38, 2003.

[22] Riichiro Mizoguchi. YAMATO: yet another more advanced top-level ontology. Proceedings of the Sixth Australasian Ontology Workshop, pages 1–16, 2010.

[23] João Moreira, Luís Ferreira Pires, Marten van Sinderen, Laura Daniele. SAREF4health: IoT standard-based ontology-driven healthcare systems. In this volume.

[24] Till Mossakowski, Mihai Codescu, Fabian Neuhaus, and Oliver Kutz. The distributed ontology, modeling and specification language–DOL. In *The Road to Universal Logic*, pages 489–520. Springer, 2015.

[25] Fabian Neuhaus, Elizabeth Florescu, Antony Galton, Michael Grüninger, Nicola Guarino, Leo Obrst, Arturo Sanchez, Amanda Vizedom, Peter Yim, and Barry Smith. Creating the ontologists of the future. 6:91–98, 01 2011.

[26] Fabian Neuhaus, Amanda Vizedom, Ken Baclawski, Mike Bennett, Mike Dean, Michael Denny, Michael Grüninger, Ali Hashemi, Terry Longstreth, Leo Obrst, et al. Towards ontology evaluation across the life cycle. *Applied Ontology*, 8(3):179–194, 2013.

[27] Adam Pease, Ian Niles, and John Li. The suggested upper merged ontology: A large ontology for the semantic web and its applications. In *Working notes of the AAAI-2002 workshop on ontologies and the semantic web*, volume 28, pages 7–10, 2002.

[28] Md Kamruzzaman Sarker, Adila Krisnadhi, David Carral, Pascal Hitzler. Rule-based OWL Modeling with ROWLTab Protege Plugin. In E. Blomqvist et al, editors, The Semantic Web. 14th International Conference, ESWC 2017, Portoroz, Slovenia, May 28 – June 1, 2017, Proceedings. Lecture Notes in Computer Science Vol. 10249, pp. 419-433. Springer, Heidelberg, 2017.