The Attractiveness of Foundational Ontologies in Industry

Stefano BORGO [a], Leonardo LESMO [b]
[a] *Laboratory for Applied Ontology (LOA), ISTC-CNR, Trento - borgo@loa-cnr.it*
[b] *Dipartimento di Informatica - Università di Torino - lesmo@di.unito.it*

Introduction

If we were to write a report on the ontology penetration rate in application-oriented domains like semantic web, database, engineering, business and medicine, just looking at the last year public events we would be justified in using an enthusiastic tone and even in going as far as to claim that ontology is nowadays a cornerstone in these areas. Indeed, many important conferences and specialized meetings devote considerable part of their time to ontology topics and are careful to register the new trends in ontological research. Prominent examples in 2007 are the OnTheMove Federated Conferences (OTM)[1] and the International Conference on Conceptual Modeling (ER)[2] in the database, business and infrastructure areas; the International Semantic Web Conference (ISWC), the Asian Semantic Web Conference (ASWC)[3] and the European Semantic Web Conference (ESWC)[4] in the domains related to Semantic Web; the International Joint Conference on Artificial Intelligence (IJCAI)[5], the Atlantic Web Intelligence Conference (AWIC)[6] and the Web Intelligence (WI)[7] in the vast Artificial Intelligence field. These are just a few of the international conferences that took place last year and explicitly related to ontology, not to mention the variety of associated workshops many of which are entirely dedicated to ontology, tools for ontologies and ontology application.

However, if we look closely at the data, we notice that these events are attended by companies and enterprises only in minimal part. Of course, one can refine this claim drawing several distinctions: between large and small/medium companies, between production industries and service providers, between developed and developing countries. Yet, there remains a feeling that the fuss about ontology is mainly at the level of research and its surrounding niches.

The Formal Ontologies Meet Industry workshop series began in 2005 to foster a positive relationship between formal ontology [1] and all the four sectors of industry: natural resources production (e.g., agriculture, fishing) and extraction (mining), manufacturing and construction, services to the business and consumers (from insurance and banking to education and health), and optimization research and design. From the experience gained in these appointments, it is clear that the gap between ontology research and industrial domain is wide and that the overlap between the two sides grows at a slow pace. We see several reasons for this, some of which have been discussed in the first FOMI report [2].

In this note we discuss another point, namely, the particular role of axiomatically rich theories based on formal semantics which are also known as *heavyweight ontologies*. In particular, among these systems we concentrate on foundational ontologies [3], that is, well designed and general heavyweight ontologies whose aim is to capture a clear perspective on reality by modeling philosophical positions (see e.g. [4]). The goal is to highlight the role of these sophisticated ontological systems in the industry domain and, vice versa, the role of industry (in the large) in ontological research.

Before entering into the discussion, it might be useful to point out that across the three FOMI editions, the majority of submissions are on the application of *lightweight ontologies*[8] to domains as far apart as chemistry, manufacturing, e-commerce, corporate knowledge, cultural heritage, network management and so on. This preference for lightweight ontologies (as opposed to heavyweight ontologies) is sometimes motivated but most of the time it is the result of a still imprecise understanding of the role of ontologies in information systems and of what one can actually do with the different ontological systems. In particular, if the layman does not always understand what distinguishes an ontological claim (any car has an engine, any car is always located somewhere) from an epistemic or factual assertion (any car has a radio device, any car can travel at 50 mph), we should not be surprised that she does not see the advantages of having part of the knowledge system deeply formalized. Similarly, we need some time before the novelties brought up by the new ontology discipline like, e.g., the formal distinction of properties, individual qualities and roles (this car *is gray* as requested, *the color of that hill* sticks out, this room *is an office*) can be fully appreciated in application fields. Naturally, the layman begins to take advantage of the new ontological perspective by exploiting it within well known techniques like, e.g., taxonomies. This attitude explains, at least in part, why today lightweight ontologies are popular in applications. Much harder is to understand the advantages of a new discipline and to exploit it in practice when the new perspective comes implemented into unfamiliar and sophisticated techniques, as in the case of heavyweight ontologies.

Lightweight ontologies, we said, do not bring new technical advantages due to their reliance on traditional approaches (classification techniques, graphical descriptions, glosses in pseudo-natural languages). Nonetheless, they witness the growing awareness of the importance of a correct terminology and of a careful (although necessarily informal) description of intended usage. The FOMI papers that are limited to lightweight ontologies are indeed quite interesting for the problems they bring up and the consequent discussions they arise.

Beside the variety of presentations based on lightweight ontologies, every FOMI event has seen some theoretical work and some example of heavyweight ontology application. Papers in the first group have been addressing the formalization of general notions, like function and product, and are helpful in forcing the practitioners and the researchers into (i) a clarification of the concepts that underlie these notion, (ii) a discussion of the variety of applications they can cover and (iii) the discovery of practical drawbacks that may be detected only working in real-world applications. More interesting is perhaps the work reported on the application of heavyweight ontologies since here the novelty of the ontological approach is seen in all its potentiality. These papers go to the core of the FOMI aims by providing information on new ontological methodologies, like the formal application of descriptive theories (e.g., mereology) to model engineering scenarios within an ontological perspective, and by showing how these can be effectively implemented. Unfortunately, up to now, papers in this perspective have been submitted by academics only. In spite of the fact that these researchers work side by side with engineers and domain experts, this observation shows how the application of heavyweight ontologies still needs the leading role of theoreticians. That is, we lack tools and explicit methodologies to put domain practitioners in the position to independently experiment these systems.

The hope is that in the next few years we will see an increasing number of this latter type of works, perhaps with stronger participation of industrial personnel. For the moment, from the debate that accompanies the presentations at the workshop, we acknowledge that the practitioners in the domains spanned by FOMI are recognizing the role of rich and structured ontologies and show interest in their potentialities. Of course, it takes more than these few papers at FOMI to push them into an active investigation of the richness of ontology research: the area has to gain in stability, clearness and maturity.

The Role of Foundational Ontologies in Industries

Nowadays a few foundational ontologies are available and are being tested in different domains and application projects. The initial hope to reach a unique general comprehensive ontology (or ontological framework) that unifies all ontological perspectives is definitely abandoned. People that see ontology as a tool to make systems or applications interoperable, find discouraging that foundational ontologies themselves suffer from the interoperability problem. If we need to rely on them – they argue – it seems we should first find a way to integrate or make them interoperate. However, this is effectively a complex issue and, notwithstanding some results, one may think that it is better to give up on foundational ontology and develop instead direct mappings between the numerous systems and modeling techniques that are in practice today. To put things straight on this issue, we have two different questions to address, namely:

- Given that there are several distinct foundational ontologies, is this a problem for interoperability?
- What is a foundational ontology good for in industry?

Regarding the first question, our answer is 'no'. The view of foundational ontology as the glue that allows us to assemble different applications into a unique coherent and interactive system is quite naïve. What these ontologies do is to formalize an explicit "view on reality" by clearly indicating what is assumed to exist and how things are assumed to relate to each other. The goal, thus, is far from having everything under one single description of reality. The idea is that, in order to reliably communicate and interact, one needs to know what others (people, agents, organizations or artificial systems) believe about reality and, it is assumed, this result is effectively achieved once one has available the foundational ontology that best captures their view of reality. That is, to be able to interact with another system, we need to have available its foundational ontology. In this way one can build a formal interface to translate information from one system to another by coding the information in the first into the view of reality that the latter adopts.

Regarding the second question, we answer with an example.

Several authors in the engineering domain have been working on the notions of function and behavior from different perspectives: artificial intelligence, system modeling, product description, and so on. B. Chandrasekaran and J.R. Josephson in [5] claim that, relatively to the areas they considered, engineers use five different meanings for the term ‘behavior’ and two for the term ‘function’ and informally discuss the relationship among these behavior(s) and function(s). To take advantage of this result, one should be able to formalize and logically relate these different meanings. This has been shown to be possible via a foundational ontology [6]. The role of foundational ontology in this work is crucial: it motivates and provides the general framework in which it becomes possible to model the different meanings of these general terms. It is important to note that in this “ontologization” the meanings given by Chandrasekaran and Josephson have been modeled without discussing their value (if engineers say these meanings are what they need, one should capture exactly and precisely those meanings) nor trying to twist them into the ontology framework (either they fit the ontology view of the world or we need a different ontology). Technically, the result is expressed in a series of logical formulas that show how these different meanings depend on an ontological notion of behavior (and function) and how they are related to each other while remaining different.

Nonetheless, one may still claim that a foundational ontology is not really necessary to reach interoperability. Perhaps, one could insist, just a set of general concepts for design and manufacturing, concepts like functionality, product, process, production plan and so on, is enough for this domain without any need to refer to a foundational system. We claim that this dismissing position underestimates the variety of the industrial domain. Indeed, a closer analysis of the case described above shows that the notions of function in Chandrasekaran and Josephson depend on the notions of behavior. However, this is not true in other approaches. The Functional Representation approach of Chandrasekaran and Josephson is just one among several that rely on these very notions like the Function-Behavior-Structure of J. G. Gero [7,8], the Function-Behavior-State of Y. Umeda [9], the Structure-Behavior-Function by A. Goel [10,11]. Instead, R. Stone and K. Wood do not even make use of the notion of behavior in their Functional Basis model [12]. In some cases the relationship between function and behavior is reversed: the very notion of behavior is seen as a specialization of that of function. For a different example, one may hope to get a shared view on specific notions like ‘shape’ or ‘electric power’, but for fairly general concepts, like operation in the shop floor, we have to face overt ontological issues if we want to be able to use such a notion together with those of plans, processes and agents’ actions.

The different approaches in the various industrial domains rely on different (often implicit) local ontologies and primitives; the problem of relating these is inherently ontological. Once we have clarified and organized the different meanings of these key concepts, we can further specialize them to coherently capture specific terminology in the different applications and representation systems.

What is Ontological Analysis Good for in Industry?

The introduction of ontological analysis [13] in artificial intelligence and knowledge modeling is motivated by several considerations: the limit of domain dependent modeling, the failure of interoperability among independently developed systems, the analysis and classification of background assumptions etc. These issues are important in the modeling of the enterprise as well as of the production process and affect the potentialities of the modern industry domain, from the exploitation of virtual enterprises to the development of integrated product lifecycles. The answer provided by the research in ontology is orthogonal to (and integrates with) the innovative techniques from other disciplines: it aims at deepening and make transparent our knowledge of the systems and of the environment they live in. For example, it clarifies the different uses of the term 'product' within the same automotive company or across the supply chain, it identifies the functionalities that are potentially realizable by services or embedded by technical artifacts, it provides uniform ways to model properties and to translate them in formats intelligible to different legacy systems.

One should not think that the aim of ontological analysis is to build good ontologies only. It helps to improve existing systems like standard databases. The ontology and the database communities yield, at a minimum, different perspectives: the purpose of ontology is to define and classify *categories* of entities (classes like *Drilling machines* as opposed to a specific instance like *The drilling machine item #123*) mainly by organizing and relating their formal properties and interactions. The purpose of databases is to collect information to describe *situations* by representing entities (classes as well as instances) and their actual relationships. The goal is the completeness (and efficient management) of the information which is deemed necessary in some domain or application, not the distinction between its ontological or factual nature. For example, an ontology must distinguish processed items from items on sale, although the two classes of entities may coincide in a given company, while a database may intentionally ignore this conceptual distinction in order to simplify the repository and improve data quality. Ontological analysis is today used to improve traditional databases as well as to create optimal and transparent interfaces for their interoperability. The role of ontologies in this latter area is, perhaps, fairly well understood. The first task consists in developing sound modeling guidelines and improving the use of well established representation languages. An example can be found in [14] where modeling constructs of attributes and datatypes are analyzed and a methodology is developed for the UML modeling language.

The Role of Industry in Foundational Ontology

Foundational ontologies are not tools for every application. For example, several applications in the Semantic Web can safely rely on lightweight ontologies because they strongly depend on statistical analysis of large sets of data, e.g., classifications based on trends or social networking, or because they explicitly give up on precision and clearness in favor of other aspects like simplicity of use, emotional descriptions, personal web, ephemeral classes and descriptions. Foundational ontologies are sophisticated and expensive to produce; although their use can be valuable in any domain (including the classification of trends and personal tagging), their construction and adoption should be motivated by a cost-benefit analysis. Domains based on sophisticated artifacts that have a relatively long life cycle and whose construction, maintenance and update requires careful analysis and considerations of aspects like functionality, requirements, implementation and sustainability (airplanes, radar systems, civil constructions, electric or oil networks, biomedical instruments and so on), find great advantages from a principled approach based on foundational ontologies. The reason lies in the real possibility, brought by these ontologies, of integrating the information across the whole industrial process: from the design phase to the after-sale services, from product update to manufacturing adaptation, functionality control and guidebook update. These observations show the interest in exploiting foundational ontologies in domains like industry and medicine beside the traditional area of the Semantic Web.

The advantages brought by foundational ontologies are easily foreseen but the deployment of these ontologies in complex domains cannot take place without some initial investment. Foundational ontologies require time to develop and in this phase in which just a few of these ontologies are available and their exploitation is just at a start, industries should invest considerable resources to build, refine or adapt these ontologies to their needs. Even more challengingly, these ontologies are too sophisticated to be used or understood by untrained people: the development of appropriate tools by which the average employee can effectively use them, perhaps relying on a training period of a few hours only, would require considerable efforts. These are real drawbacks and should be openly faced in order to understand the real industrial needs and to address investments in the right direction. Right now, most projects that rely on ontology in this area are still based on public funding and only in some rare cases are supported or initiated by industrial consortia (the example here is the EPISTLE Core Model[9], an ISO standard which has been recently proposed as a top-level ontology).

Since foundational ontologies are, after all, new tools for industrial needs, the development of an ontology as an open or proprietary standard depends on industrial considerations. However, if we want to take advantage of web technology and the new approaches toward virtual enterprises and integrated supply chains (just to name a few cases), we see that these general ontologies will have major impact and will provide the biggest advantages if they are publicly shared and widely adopted, or at least widely recognized: due to the costs of restructuring industrial information systems, many enterprises may at first prefer to enhance their legacy or proprietary systems by providing an interface that aligns the enterprise data and knowledge structures to a standard ontology. In this way, an enterprise can take advantage of a standard ontology from the beginning and avoid to redesign at once the whole information system to conform to the ontology: a change that requires investment in terms of money, time, and personnel training.

It is sometime claimed that foundational ontologies are not suited if the target domain is quickly evolving as it happens, e.g., in software and artifacts based on new technologies. The argument, as far as we see, relies on the confusion between types of ontologies (foundational, core, domain, formal, lightweight, upper level etc.) and knowledge bases (or even databases). Foundational ontology sets the knowledge structure by establishing the meaning of the concepts central to a given domain, thus it *defines* the very domain at stake: in the industry domain a foundational ontology would be extended to formally represent concepts like artifact, component, feature, function, process, service, operation, agent, and so on. The evolving set of products and product models or the specific functionality of an item are pieces of information that one finds in knowledge bases, not in foundational ontologies, and knowledge bases have long proved that they can cope with evolving environments, provided they are well constructed and maintained. Ontologies can help in providing the correct framework on top of which to construct optimal knowledge bases and databases for the domain perspectives and needs, which includes the capability to model new product types and to construct and discover new functionality types. In principle, a foundational ontology is quite stable over time if we exclude possible extensions to include new general concepts.[10] Companies should learn to distinguish the different types of ontologies since these are developed to answer different types of problems. In particular, they should be able to distinguish between foundational ontologies and knowledge bases; two complementary systems whose alignment is crucial for the success of the evolution toward the ontology-based enterprise.

Finally, industry should be less shy in addressing the research community regarding ontology and ontology applications. The gap between ontology research and ontology implementation has brought many researchers to spend most of their efforts toward ontology languages and reasoning classifications. The consequence is that crucial aspects to improve usability [15] like terminology development (needed to foster understanding and correct usage by non-experts) and ontology interfaces for the end-user [16], do not hold the stage today. Terminology development and ontology interface development are of course just two aspects of the pervasive relationship between ontology, natural language and human/computer interaction, a relationship which today has to be assessed with respect to the main source of information: the Web.

Ontology, Language and Communication

It is well known that the vast majority of information sources is based on natural language. Most Web pages include portions of text written in one of the hundreds of languages of the world. The next step in the process of making Web information available to an even larger set of users is to provide for access in written or spoken language, possibly exploiting the current technology for dialogue management. But this can hardly be achieved without an in-depth shared view on the meaning of the individual words.[11]

Unfortunately, also in this field, the huge number of technical and scientific papers describing possible relations between language and ontologies (ranging from the use of ontologies for understanding texts to the use of texts for building or extending ontologies) is not mirrored in a comparable number of real-world applications. Nonetheless, some software products for specific fields of application do exist; the most relevant example concerns perhaps healthcare (`http://www.landcglobal.com/`). The limited diffusion of NLP tools exploiting (lightweight) ontological knowledge can be taken as supporting two claims: ontologies are useful for language analysis, but their development is highly expensive (requiring consistent fundings, as are more easily available in the medical field).

The first point is hard to deny: the extraction of the meaning of a piece of text involves the representation of this meaning in some suitable "internal language" enabling a machine to perform various types of reasoning. Of course, this is not needed if what is needed is a link to a web page (but the current techniques for Web lookup can hardly be said to involve meaning extraction) or a summary (where the representation language is the same as the original text). But if an internal language is needed, it must have some *formal* flavour; usually, it will be a logic language, including predicates and constants.[12] If we want to move toward the sharing of this *meaning* among systems, there must be an agreement about the predicates used and about what they mean. This can be achieved (at least in part) just via an agreement on the meaning of terms, as the one provided by the adoption of shared lightweight ontologies.

The second point, i.e. the high cost incurred in the development of ontologies, is not less agreed upon. The obvious question that an industry or a public administration asks itself is: is it worth? Though this question and the possible answers have been analyzed in depth in the previous sections, something should be added here. Much effort is being devoted to the automatic development of ontologies, especially lightweight ones. Usually, the input data for this development are pieces of text, but the current status of these automatic techniques yield results that are only partially encouraging. Again, it seems that there is no way out, especially in the case one wants to face the problems associated with the "ephemerality" of classes mentioned in the previous section. Arguably, more effort is required on this topic, possibly adopting approaches (that are based on a sort of *bootstrapping*) where language is analysed in more depth, on the basis of established ontologies where the meaning of terms is reasonably fixed and shared. However, it is hard to imagine that these efforts can achieve their goal in a short time without a strict cooperation between industry and academy.

But language also teaches us a useful lesson: if I receive a "Call for participation to the FOMI workshop", I am usually able to understand what it means, i.e. what is a *workshop*, what would involve for me to *participate*, what the *call* has been sent for. The same probably holds for my friends and colleagues in Japan, Germany, India, South Africa, etc.. Concepts as *workshop* (a complex event) or *participation* (perhaps an activity) are notoriously difficult to model in a heavyweight ontology and receive different formalizations in different approaches. Nonetheless, people growing up in different cultures, speaking different languages and having different life experiences, are successful to talk to each other and to *understand* each other. One may ask: Are they using the same or different ontologies? Is there a common core that is *common* in the true sense of this word? Can some form of ontology capture this core? What combination of semantic interaction components should we look for?

Having acknowledged in this paper the existence and value of different ontological approaches and their different goals, let us close with a word of hope: maybe not in the next, but in some FOMI workshop close in the future, we will discuss about different ontologies acting as the basis for real interoperation: interoperation between humans and humans (via web pages written in different languages), between humans and machines (where we will be able to give commands to robots in such a way that we and the robot have a common understanding about what a command or a required operation is), and interoperation between computer systems. At that point, we will be able to say that the final goal of ontological studies for improving everyday life and industrial applications is being approached.

Notes

[1] `http://www.cs.rmit.edu.au/fedconf/`
[2] `http://er2007.massey.ac.nz/`

[3] http://iswc2007.semanticweb.org/main/default.asp

[4] http://www.eswc2007.org/programmecommittee.cfm

[5] http://www.ijcai-07.org/

[6] http://www.awic2007.net/call-for-papers

[7] http://www.cs.sjsu.edu/wi07/

[8] The term is generally applied informally. Here we use it to refer to semantically weak systems like those based on taxonomies, concept maps, or conceptual schemata.

[9] See the "Industrial automation systems and integration – Integration of life-cycle data for process plants including oil and gas production facilities – Part 2: Data model"; it can be accessed at: http://www.tc184-sc4.org/wg3ndocs/wg3n1328/lifecycle_integration_schema.html.

[10] Think, e.g., about the new concepts brought into the manufacturing based economy by the globalization and information technology.

[11] "In order to perform the kind of reasoning/inference required for deeper (semantic) understanding of texts, as required for high-quality Machine Translation, Summarization, and Information Retrieval, it is imperative to provide systems with a wide-ranging semantic thesaurus."(Objectives of the SENSUS Project http://www.isi.edu/natural-language/projects/ONTOLOGIES.html)

[12] If one gets a DB query, then it will include names of relations and fields and values appearing therein.

References

[1] N. 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.

[2] Stefano Borgo, Matteo Cristani, and Roberta Cuel. Formal Ontology meets Industry. *Applied Ontology*, 1(3-4):217–220, 2006.

[3] Stefano Borgo. How Formal Ontology can help Civil Engineers. In Jacques Teller, John Lee, and Catherine Roussey, editors, *Ontologies for Urban Development*, Studies in Computational Intelligence, pages 37–45. Springer, 2007.

[4] C. Masolo, A. Gangemi, N. Guarino, A. Oltramari, and L. Schneider. Wonderweb deliverable d17: The wonderweb library of foundational ontologies. Technical report, 2002.

[5] B. Chandrasekaran and J.R. Josephson. Function in device representation. *Engineering with Computers*, 16(3/4):162–177, 2000.

[6] Stefano Borgo, Massimiliano Carrara, Pawel Garbacz, and Pieter E. Vermaas. A Formal Ontological Perspective on the Behaviors and Functions of Technical Artifacts. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 23(1):(to appear), 2009.

[7] J.S. Gero. A knowledge representation schema for design. *AI Magazine*, 11(4):26–36, 1990.

[8] J.S. Gero and U. Kannengiesser. The situated Function-Behavior-Structure framework. *Design Studies*, 25:373–391, 2004.

[9] Y. Umeda, M. Ishii, M. Yoshioka, Y. Shimomura, and T. Tomiyama. Supporting conceptual design based on the function-behavior-state modeler. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 10:275–288, 1996.

[10] A. Goel. A model-based approach to case adaptation. In *Proc. 13 Ann. Cogn. Sci. Conference*, pages 143–148. Mahwah, NJ: Lawrence Erlbaum, August 1991.

[11] S. Bhatta and A. Goel. Design patterns and creative design. pages 271–284, Berlin, 2002. Springer.

[12] R.B. Stone and K. Wood. Development of a functional basis for design. *Journal of Mechanical Design*, 122(4):359–276, 2000.

[13] Nicola Guarino. Formal ontology, conceptual analysis and knowledge representation. *International Journal of Human and Computer Studies*, 43(5/6):625–640, 1995.

[14] Giancarlo Guizzardi, Claudio Masolo, and Stefano Borgo. In defence of a trope-based ontology for conceptual modeling: An example with the foundations of attributes, weak entities an datatypes. In *Conceptual Modeling - ER'06*, volume LNCS 4215, pages 112–125, Tucson, Arizona (USA), 2006. Springer.

[15] Mark S. Fox and Michael Gruninger. On ontologies and enterprise modeling. In *Proc. Int'l Conf. Enterprise Integration Modeling Technology*. Springer-Verlag, 1997.

[16] Martin Hepp. Possible ontologies: How reality constrains the development of relevant ontologies. *IEEE Internet Computing*, 11(1):90–96, 2007.