

Ontology Summit 2010 – Creating the Ontologists of the Future

Fabian Neuhaus, Elizabeth Florescu, Antony Galton, Michael Gruninger, Nicola Guarino, Leo Obrst, Arturo Sanchez, Amanda Vizedom, Peter Yim, and Barry Smith

Abstract. The goal of the Ontology Summit 2010 was to address the current shortage of persons with ontology expertise by developing a strategy for the education of ontologists. To achieve this goal we studied how ontologists are currently trained, the requirements identified by organizations that hire ontologists, and developments that might impact the training of ontologists in the future. We developed recommendations for the body of knowledge that should be taught and the skills that should be developed by future ontologists; these recommendations are intended as guidelines for institutions and organizations that may consider establishing a program for training ontologists. Further, we recommend a number of specific actions for the community to pursue.

Keywords: Education, Ontology Summit

Editorial Remarks

This document is the Joint Communiqué of the Ontology Summit 2010, the fifth in an annual sequence of events in which the international ontology community explores a given topic over a series of online meetings culminating in a face-to-face symposium held at the National Institute of Standards and Technology (NIST). The goal is to reach a position enjoying a broad consensus across the entire community. The 2010 Ontology Summit was devoted to the education of ontologists under the heading ‘Creating the Ontologists of the Future’. It was organized by the Ontolog Forum, NIST, the National Center for Ontological Research (NCOR), the National Center for Biomedical Ontology (NCBO), and the International Association for Ontology and its Applications (IAOA). Our goal was to get an overview of how ontologists are currently trained, whether the training meets the requirements that ontologists face during their work, and of the degree to which this training is sufficient to meet projected needs. During roughly three months the participants of the Ontology Summit organized seven teleconferences, two surveys, and a Delphi study. The process culminated in a face-to-face meeting in Gaithersburg, Maryland on March 15-16th, 2010, where this document was finalized. Its purpose is to communicate to the larger ontology community and to the general public how the current training of ontologists falls short and what steps should be taken to improve the situation. The Ontology Summit 2010 was co-sponsored by 61 organizations and this Joint Communiqué was endorsed by 93 members of the extended ontology community.

1. Summary

Increasingly, major national and international projects and systems centered on ontology technology are being developed and deployed by governments and by scientific and commercial organizations. This brings a growing need for ontology expertise and thus for new methods and organizations for the education and training of ontologists. The goal of the Ontology Summit 2010 was to develop a strategy for the education of ontologists. To achieve this goal we studied how ontologists are currently trained, the requirements identified by organizations that hire ontologists, and developments that might impact the training of ontologists in the future.

The main findings and results of the Ontology Summit 2010 are:

1. that there is already a large demand for trained ontologists, and the demand is expected to increase as ontology-based technologies become more successful and as the quantities and number of different types of data continues to expand.
2. that there are very few formal training opportunities for ontologists, and they often do not meet the needs of trainees or of those who would hire them.
3. that organizations wishing to hire ontologists often have difficulties in identifying qualified candidates since there are so few formal qualifications in ontology, and there is no professional organization that certifies ontologists.

We developed recommendations for the body of knowledge that should be taught and the skills that should be developed by future ontologists; these recommendations are intended as guidelines for institutions and organizations that may consider establishing a program for training ontologists. Further, we recommend a number of specific actions for the community to pursue that will improve the education of ontologists.

2. Introduction

Currently, data and information are often siloed, reflecting the fact that they have been collected in ways designed to address narrowly tailored local needs and in the context of specific applications. As a result, data is difficult to reuse for new purposes; different bodies of data do not cumulate; and possible benefits of data integration are lost.

Applied ontology is designed to counteract these effects by creating so-called ‘ontologies’ that are designed to facilitate more effective information exchange through machine-interpretable representations of reality of more global validity and scope. To this end, applied ontologists develop the theories, methods and formal tools to support the creation, use and evaluation of ontologies.

Ontologies play a central role in the Semantic Web, the Linked Data movement, and in many other technological developments, for example in the areas of semantic services and the semantic enterprise. Multiple ontology-based approaches, loosely grouped under the heading ‘semantic interoperability’, have come to the fore as potential solutions to critical interoperability problems. Further, technologies that incorporate and rely on ontologies are used to increase transparency both within and across organizations, and also to enhance communication not only between computers but also between human beings.

Major national and international ontology projects have been initiated by governmental, scientific and industrial organizations, for example to support exchange of information across scientific, organizational or linguistic boundaries. But the success of such efforts depends on the availability of well-trained ontologists, capable of designing and building the needed representations and of supporting their successful implementation in the integration of data and information.

It is already clear that the resultant need for persons with ontology expertise goes far beyond the current availability of appropriately trained personnel. Organizations seeking to hire ontologists often face difficulties in identifying qualified candidates since there is no professional organization that certifies ontologists and few educational institutions that offer formal education and training in ontology.

Enhanced training would at the same time provide a developing body of knowledge not only concerning the techniques of ontology but also concerning important successes and failures. In this way, it would help those working in semantic technology and related fields to recognize where ontology can be successfully used, and at the same time to avoid a variety of characteristic errors – and resultant project failures – that have affected ontology initiatives in recent years.

To work effectively, the ontologist must command a specific set of skills, and it is important to examine how formal education and training can help both to meet the increasing demand for those who have these skills, and to enable project managers to distinguish qualified ontologists from those who simply claim the title.

The goal of the 2010 Ontology Summit was to develop a strategy for a more coherent approach to the education and training of ontologists, building on the results of previous Ontology Summits

(<http://ontolog.cim3.net/cgi-bin/wiki.pl?OntologySummit>). To achieve our goals the community conducted two surveys, a Delphi study and several panel discussions, to address the following questions:

1. How are ontologists currently trained?
2. What abilities do ontologists consider as necessary for their work?
3. What do employers expect from individuals that are hired as ontologists?
4. What are the developments that might impact the training of ontologists in the future?

The responses to these and related questions allowed us to identify a number of different career paths for ontologists as well as the associated knowledge and skills. On this foundation we developed recommendations for the content that should be taught to future ontologists. In the following we will present the results of our findings as well as our recommendations.

3. Current State of Ontologist Training

In this section we summarize our key findings about the current state of the training of ontologists, requirements for training opportunities, and expected developments. More details can be found at the Ontology Summit 2010 website (<http://ontolog.cim3.net/cgi-bin/wiki.pl?OntologySummit2010>).

The demand for ontologists is expected to rise considerably. There is general consensus from the panel of experts that ontology-enabled applications and tools will become ever more widely used in the coming years, and that a correspondingly large number of people with adequate ontology education and training will be needed; about 5% of information systems and software engineering professionals in the next 5-10 years.

There is a large gap between education needs and education availability. Based on the surveys we conducted, only one academic program was identified as devoted to education in applied ontology (a master's program at the University of Buffalo). In addition, we identified 21 programs that offer ontology-centered courses. These are typically masters-level courses that form part of computing programs. The institutions that have been identified as offering at least one course with partial coverage of ontology are located in Belgium, Brazil, Germany, Iran, Italy, Japan, Netherlands, UK, and the USA. These results show that some students (mainly in computer science programs) have some exposure to ontology; and that interest in ontology is not restricted to a particular geographic region. However, the main result is that there exists only one academic program devoted to the training of ontologists. As a result, most of those who might consider an ontology career currently have formal training in other fields and must therefore resort to on-the-job or self-directed training in ontology.

Significant demand for training opportunities for working professionals. We found that most training opportunities exist within academic degree programs – for example, ontology courses within computer science curricula. However, there is substantial demand for training outside of such contexts, including single courses, professional certification programs, hands-on training with or without certification, and familiarization courses as short as one week. This demand is perhaps best interpreted in conjunction with the fact that the majority of respondents who expressed an interest in training already have some level of familiarity with ontology and indicated that their interest in the field grows out of their current work.

Available training opportunities for professionals do not meet needs. Logic and formal semantics are identified as a requirement for ontologists by potential employers, evaluators, working ontologists, and potential trainees. Typically, these subjects are covered in academic programs; however, they are not included in shorter-duration training programs offered to working professionals. There is therefore a significant need for more substantial training programs, delivering not just familiarity but also technical competence, and offered in a way that makes them accessible also to those not pursuing an academic degree.

<i>Abilities required for developing, improving, and applying ontologies:</i>	
1. Clarifying the purpose of a given ontology, understanding potential deployment, performing requirements analysis	sources; relevant data; domain expertise, ontology expertise)
2. Analyzing existing legacy models and data that are relevant to a given project	9. Assembling an ontology from reusable modules
3. Judging what kinds of ontologies are useful for a given problem (including: knowing when ontologies are not useful)	10. Using (reading, writing) different representation languages
4. Managing ontologies across their life cycle (requirements analysis and planning, managing a systematic update process, versioning, documentation, help desk, ...)	11. Conducting ontological analysis, that is identifying entities and relationships; formulating definitions and axioms
5. Identifying, evaluating and using software tools that support ontology development	12. Evaluating and improving ontologies (finding errors via manual term-by-term inspection, solving interoperability problems, decomposing large ontologies into interconnected modules)
6. Choosing the appropriate representation language	13. Documenting ontologies (e.g., providing natural language definitions and providing concise explanations for axioms)
7. Choosing the appropriate level of detail	14. Working in teams, including those which support the distributed development of ontologies
8. Identifying existing content resources (e.g., existing ontologies, terminologies and related re-	15. Using at least one modern programming/scripting language

Fig. 1. Core Skills

Important subjects are absent from existing curricula. Experienced, working ontologists, potential trainees, potential employers, professional collaborators, and senior professionals who manage, lead, evaluate and depend upon ontologists identified a number of important subjects that are largely absent from existing curricula.

Ontology is interdisciplinary. The surveyed ontology experts expect ontology education to take place in interdisciplinary programs. Further, they themselves come from a wide variety of backgrounds, and they consider their non-ontology training as relevant for their ontology-related work. These results suggest that there is much to be gained by designing curricula that incorporate contributions from multiple disciplines and welcome students from multiple backgrounds, in contrast to focusing more exclusively on just a few feeder fields or departments.

Employers cannot easily recognize qualified ontologists. Because of the small number of designated programs for ontologists, most people working in the field have no formal qualification in ontology. Further, there does not exist a professional organization that certifies ontologists. Because of the lack of formal qualifications, institutions that intend to hire ontologists often have difficulties in identifying qualified candidates.

4. Recommendations for the training of ontologists

Based on our findings we present lists of the topic areas in which a student should be taught and the abilities that should be developed in an ontology program. In Tables 1-3 we distinguish between *skills* (the ability of a student to do something) and *knowledge* (basic notions grasped). Since skills build on knowledge, the two must be taught together.

Since ontology is a highly interdisciplinary field and because the careers of ontologists are diverse, it is unrealistic to expect students to learn everything that might be relevant. For this reason, one could charac-

- | | |
|---|---|
| <ol style="list-style-type: none"> 1. The basic terminology of ontology (relation of ontology to knowledge representation, conceptual modeling, data modeling, ...) 2. Theoretical foundations <ol style="list-style-type: none"> (a) first-order logic, basics of description logic, modal logic, and second-order logic (b) set theory (c) basic notions of philosophical ontology (universals and particulars, mereology, essence and identity, unity and plurality, dependence, change in time, ...) (d) philosophy of language (the use-mention confusion, sense and reference, speech act theory, ...) (e) knowledge representation, conceptual modeling, data modeling; metadata 3. Representation languages Part 1: Resource Description Framework (RDF), Web Ontology Language (OWL); Common Logic 4. Building and editing ontologies <ol style="list-style-type: none"> (a) human aspects (application of classification principles, manual auditing, ...) (b) software tools (Protégé, ...) (c) addressing interoperability problems among ontologies 5. Ontology evaluation strategies and theories (Ontoclean, ...) | <ol style="list-style-type: none"> 6. Examples of ontologies, illustrating different methodologies <ol style="list-style-type: none"> (a) upper-level ontologies (Basic Formal Ontology (BFO), Descriptive Ontology for Linguists and Cognitive Engineering (DOLCE), Suggested Upper Merged Ontology (SUMO), ...) (b) mid-level, domain-spanning ontologies (Process Specification Language, ...) (c) domain ontologies (Gene Ontology, Enterprise Ontology, ...) 7. Examples of ontology applications (successes and failures) <ol style="list-style-type: none"> (a) as controlled vocabularies / standards, to achieve coordination among human beings (b) to solve interoperability problems among external data resources (c) reasoning with ontology content (d) improving search and retrieval (e) natural language processing (f) decision support, situational awareness, information fusion, anomaly detection 8. Ontology and the Web <ol style="list-style-type: none"> (a) general foundations (URIs, XML, etc) (b) Semantic Web initiative (c) semantically enhanced publishing, literature annotation, data curation |
|---|---|

Fig. 2. Core Knowledge

terize our task as being one of identifying the most important knowledge and skills that an ontologist needs to do his or her job. Thus, we distinguish in Figures 1-3 between core and elective skills and knowledge. The idea is that any student should be required to gain competence in all of the core areas and some of the elective skills and knowledge. How this content should be taught is beyond the scope of this document: this is something that needs to be decided by each individual educational institution on the basis of available resources. At least some of the content is likely to be covered by existing courses in other programs (e.g., in computer science, information studies, or philosophy). We stress, however, that benefits accrue from the maximal possible degree of coordination and of shared content between those offering ontology training programs.

One challenge in creating recommendations for the education of ontologists is that ontology is a young discipline and thus has as yet no widely agreed upon body of shared knowledge, established methodologies, or common terminology. Instead, multiple terminologies are used in the different subfields of ontology, for example, deriving from specific programming environments, from database design and the conceptual modeling community, or from traditional philosophical ontology. This is a large obstacle for communication between ontologists and the users of ontologies, and we strongly recommend that all ontologist training programs include survey modules designed to familiarize trainees with these multiple terminologies.

<p>Elective Skills</p> <ol style="list-style-type: none"> 1. Coordinating ontology development efforts 2. Creating meaningful visualizations of ontology structure for use by human beings 3. Training people in the use of ontologies 	<ol style="list-style-type: none"> (b) natural language processing, natural language generation (c) cognitive theories of categorization <p><i>Supporting tools, technologies and methodologies</i></p> <ol style="list-style-type: none"> 1. Representation languages Part 2 (Semantic Web Rule Language (SWRL), Rule Interchange Format (RIF), Simple Knowledge Organization System (SKOS); OBO Format; Unified Modeling Language (UML); Entity Relationship Diagrams, IKRIS Knowledge Language (IKL), ...) 2. Ontology content acquisition (role of text mining, ...) 3. Achieving ontology interoperability 4. Principles for building ontology repositories 5. Usability and user interface issues (visualization / usability, principles of meaningful arrangement, ...)
<p>Elective Knowledge <i>Underlying and related disciplines</i></p> <ol style="list-style-type: none"> 1. Advanced logic (modal logic, temporal logic, default logic, ...) 2. Advanced philosophical ontology (mereotopology, tropes, ...) 3. Computer science <ol style="list-style-type: none"> (a) formal languages, formal machines, computability, complexity (b) automated reasoning (c) database theory (d) artificial intelligence (e) logic programming 4. Linguistics / cognitive sciences <ol style="list-style-type: none"> (a) distinction between syntax, semantics, and pragmatics 	<p><i>Application domains</i></p> <p>Any domain can be an application domain for ontologists. Ontologies are already used in many domains, including science, medicine, business, government, military, education and culture.</p>

Fig. 3. Elective Knowledge and Skills

Another challenge is that there are many diverse career opportunities open to ontologists. In the following we illustrate this diversity and the resulting differences in training requirements with two examples. *IT-oriented ontologists* are actively engaged in the deployment of IT systems involving many components in addition to the ontology itself. For these ontologists it is essential to know how to integrate the ontology into the associated applications. For this purpose ontologists need some background in software engineering, information systems design, system development, object-oriented programming, and data analysis. *Community-oriented ontologists* specialize in developing ontologies within a given domain in collaboration with subject-matter experts in relevant areas. One of the main tasks of such ontologists is to facilitate the resolution of ambiguities in such a way as to build consensus within the different communities involved. To fulfill this role, ontologists need not only to know the scientific areas covered by the ontologies (e.g., protein biology or infectious disease), but also to possess the human-oriented skills that enable them to lead teams of domain experts or to build communities that will support the effective use of ontology resources.

The core knowledge and skills that we list in Figures 1 and 2 cover the basics any ontologist will need. They are not of themselves sufficient to support a career as an ontologist; this will require either some additional background in systems development or domain-specific knowledge in some relevant application domain.

There is a strong consensus within the community that although much academic knowledge is relevant for ontologists, many important skills cannot be learned from lectures alone. Any education of ontologists has to involve hands-on training in the development and application of ontologies. Ideally, academic programs should offer their students the opportunity to gain some of this experience by participating in projects that apply ontologies to the solution of non-trivial problems in real-world environments.

5. Towards Better Education and Training of Ontologists

This document identifies the skills and knowledge a student should possess after successfully completing an ontology program. These recommendations are based on extensive studies of the current training situation, the requirements ontologists face, and the developments that may impact the situation of ontologists in the future.

To improve the training situation in applied ontology we recommend the following actions:

- Each item on the list of requisite knowledge and skills for ontologists provided above should be described in more detail.
- The development of a registry, allowing members of the community to add information about ontology-centered educational and training initiatives.
- Inclusion of more ontology-related content into model curricula for computer science (e.g., those of ACM/IEEE <http://www.acm.org/education/curricula-recommendations>) and associated disciplines.
- The creation of a wiki to collect descriptions of case studies demonstrating the importance of certain ontology engineering decisions. These might include examples of bad decisions, the problems they caused, the associated costs, and how the problems were corrected.

The requirements survey revealed a non-alignment between the training available to ontologists and the kind of training they need. As the field of ontology continues to evolve and training demand shifts in tandem, we recommend conducting similar surveys at regular intervals. This will enable training providers to ensure that their courses meet the needs of their students.

One important conclusion is that after more than two decades applied ontology still has no accepted body of shared knowledge, techniques, and criteria for evaluation. It is in part for this reason that so few ontology training programs in universities have been developed. We recommend taking advantage of the need for trained ontologists, and thus for improved ontology training, as an argument for investing effort in establishing the requisite shared body of knowledge.

Most importantly, we strongly encourage educational institutions to establish programs that address the growing need for ontologists based on the guidelines set forth in this document.

6. Endorsement

The above Communiqué has been endorsed by the individuals listed below. (Please note that these people made their endorsements as individuals and not as representatives of the organizations they are affiliated with.)

Sameera Abar	Carola Eschenbach	Finith Jernigan	Deborah MacPherson	Ravi Sharma
Jeffrey Abbott	Elizabeth Florescu	Ivan Jureta	Howard Mason	Barry Smith
Mauricio Almeida	Andrew Frank	Vicky Karen	Riichiro Mizoguchi	Bob Smith
Ken Baclawski	Alexandra Galatescu	Tomi Kauppinen	Jaimie Murdock	Jerry Smith
John Bateman	Antony Galton	Maria Keet	Mark Musen	Simon Spero
Denise Bedford	Bart Gajderowicz	Pavithra Kenjige	Fabian Neuhaus	Dagobert Soergel
Joel Bender	Jerry Glenn	Mitch Kokar	Pete Nielsen	John Sowa
Mike Bennett	Ted Gordon	Martin Kuiper	Leo Obrst	Sharon Stern
Gary Berg-Cross	Pierre Grenon	Werner Kuhn	Frank Olken	George Strawn
Mike Bergman	Michael Gruninger	Oliver Kutz	Jack Park	Vlad Tanasescu
Julita Bermejo Alonso	Nicola Guarino	Minna Lanz	Carlo Penco	Carlos Toro
Olivier Bodenreider	Giancarlo Guizzardi	Ken Laskey	Gerald Radack	Michael Uschold
Stefano Borgo	Ali Hashemi	Eric Lindahl	Steven Ray	Laure Vieu
Bruce Bray	Pat Hayes	Frank Loebe	Michelle Raymond	Amanda Vizedom
Rex Brooks	Matthew Hettinger	Terry Longstreth	Michael Riben	Matthew West
Peter Brown	Bill Hogan	Joanne Luciano	Cameron Ross	Nancy Wiegand
Patrick Cassidy	Doug Holmes	Cecil Lynch	Alan Ruttenberg	Peter Yim
Jim Disbrow	Krzysztof Janowicz	Maria Luiza Machado	Arturo Sanchez	Gian Piero Zarri
Ed Dodds	Ludger Jansen	Campos	Todd Schneider	