

Abstract

This introductory course, “Ontology engineering for the Semantic Web”, covers ontologies and semantic web in a comprehensive manner; it presents attendants with basic concepts fully illustrated with practical examples. Methodologies and tools are presented together with examples of data sources that use ontologies. Use cases will be drawn from the Biomedical Domain as well as from the Social Web. Ontologies such as VIVO (Social Web for Research and Researchers), OBO, FMA, Smart Protocols (biomedical domain) will be presented with data sources to be queried over SPARL endpoints.

Aims of the course:

Purpose of the course is to introduce students to theoretical foundations and principles of ontology engineering and also to provide overview of software tools and techniques for development and using ontologies. In addition attendants will also be presented with practical examples of Semantic Web technology; participants should be able to understand modeling and technology issues. This course is thought and designed to provide a practical overview of ontology engineering and semantic web technology.

This course is designed to be a five hours tutorial; there are no pre-requisites. This is not a hands-on tutorial. The limitation in the number of attendants is related to the space available in the venue.

Acknowledgments

I have “taken” from many sources. For instance, Oscar Corho, Maria Keet, Pedro Szekely, and others. My own experience when explaining why to develop ontologies and why should you consider semantic technologies has also impacted the design of this course.

Course contents

Introduction to Ontology Engineering, with emphasis on Semantic Web Technologies

1. Introduction

- 1.1 What is an ontology?
- 1.2 What is the usefulness of an ontology?
- 1.3 What is the Semantic Web?
- 1.4 Why the Semantic Web?
- 1.5 Success stories
 - 1.5.1 Ontology Engineering –Biomedical Ontologies
 - 1.5.2 Semantic Web -VIVO

2. Ontology and conceptual modeling, languages and tools

- 2.1 OWL
- 2.3 OWL 2
- 2.4 RDF
- 2.5 Concept Mapping
- 2.6 Protégé and Bioportal

3. Methods and methodologies

- 3.1 Design parameters and their dependencies
- 3.2 Methods: OntoClean and Debugging
- 5.3 Methodologies and tools, engineering ontologies

4. Foundational and Top-down aspects

- 4.1 Foundational ontologies
- 4.2 Part-whole relations
- 4.3 Ontology Design Patterns

5. Bottom-up ontology development

- 5.1 Relational databases and related 'legacy' KR
- 5.2 Natural language processing and ontologies
- 5.3 Biological models and thesauri

6. Querying the Semantic Web

- 6.2 Useful sources
 - 6.2.1 DBPEDIA
 - 6.2.2 Bio2RDF
 - 6.2.3 VIVO
- 6.1 SPARQL
- 6.3 Semantic Web in action –VIVO mash-ups

Recommended Reading

- Ontologies

Steffen Staab and Rudi Studer (Eds.). Handbook on ontologies. 2009.

Zalta (Ed.). Stanford Encyclopedia of Philosophy. 2010.

- Ontology Engineering

Gomez-Perez, A., Fernandez-Lopez, M., Corcho, O. Ontological Engineering. Springer Verlag London Ltd. 2004.

- Ontologies in specific subject domains

Baker, C.J.O., Cheung, H. (Eds). Semantic Web: revolutionizing knowledge discovery in the life sciences. Springer: New York, 2007, 225-248.

- Semantic Web Technologies

Pascal Hitzler, Markus Kroetzsch, Sebastian Rudolph. Foundations of Semantic Web Technologies. Chapman & Hall/CRC, 2009, 455p.

- Various subtopics in ontologies

Jerome Euzenat and Pavel Shvaiko. Ontology Matching. Springer. 2007.

Heiner Stuckenschmidt, Christine Parent, Stefano Spaccapietra (Eds.). Modular Ontologies—Concepts, Theories and Techniques for Knowledge Modularization. Springer. 2009.

Chu-ren Huang, Nicoletta Calzolari, Aldo Gangemi, Alessandro Lenci, Alessandro Oltramari, Laurent Prevot (Eds.). Ontology and the lexicon. Cambridge University Press. 2010.

- General background material

Frank van Harmelen, Vladimir Lifschitz and Bruce Porter (Eds.). Handbook of Knowledge Representation. Elsevier, 2008, 1034p.

Hedman, S. (2004). A first course in logic—an introduction to model theory, proof theory, computability, and complexity. Oxford: Oxford University Press.

F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, and P. F. Patel-Schneider (Eds). The Description Logics Handbook – Theory and Applications. Cambridge University Press, 2003.

J. E. Hopcroft, R. Motwani, and J. D. Ullman. Introduction to Automata Theory, Languages, and Computation. Pearson Education, 2nd ed., 2001.

Halpin, T., Morgan, T. Information modeling and relational databases. 2nd Ed. Morgan Kaufmann. 2008.

- Selection of journals that publish papers about ontologies

- Applied Ontology
- Journal of Web Semantics
- Journal on Data Semantics
- Artificial Intelligence
- Journal of Automated Reasoning
- International Journal of Metadata, Semantics and Ontologies

- Recommended Bibliography

- [1] Guarino, N. Formal Ontology in Information Systems. Proceedings of FOIS'98, Trento, Italy, June 6-8, 1998. IOS Press, Amsterdam, pp. 3-15
- [2] Nicola Guarino. The Ontological Level: Revisiting 30 Years of Knowledge Representation. In: A.T. Borgida et al. (Eds.), Mylopoulos Festschrift. Springer LNCS 5600, 52–67. 2009.
- [3] Nicola Guarino, Daniel Oberle, and Steffen Staab. What Is An Ontology? In: S. Staab and R. Studer, Handbook on Ontologies, Chapter 6. Springer. 2009. pp1-17.
- [4] The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. Nature Genetics, May 2000;25(1):25-9.
- [5] Barry Smith, Michael Ashburner, Cornelius Rosse, Jonathan Bard, William Bug, Werner Ceusters, Louis J. Goldberg, Karen Eilbeck, Amelia Ireland, Christopher J Mungall, The OBI Consortium, Neocles Leontis, Philippe Rocca-Serra, Alan Ruttenberg, Susanna-Assunta Sansone, Richard H Scheuermann, Nigam Shah, Patricia L. Whetzel, Suzanna Lewis. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. Nature Biotechnology 25, 1251-1255 (2007).
- [6] Berners-Lee, Tim; James Hendler and Ora Lassila (May 17, 2001). The Semantic Web. Scientific American Magazine.
- [7] Joshua S. Madin, Shawn Bowers, Mark P. Schildhauer and Matthew B. Jones. (2008). Advancing ecological research with ontologies. Trends in Ecology & Evolution, 23(3): 159-168
- [8] Wolstencroft, K., Stevens, R., Haarslev, V. Applying OWL reasoning to genomic data. In: Semantic Web: revolutionizing knowledge discovery in the life sciences, Baker, C.J.O., Cheung, H. (eds), Springer: New York, 2007, 225-248.
- [9] Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, Antonella Poggi, Mariano Rodriguez-Muro, and Riccardo Rosati. Ontologies and databases: The DL-Lite approach. In Sergio Tessaris and Enrico Franconi, editors, Semantic Technologies for Information Systems - 5th Int. Reasoning Web Summer School (RW 2009), volume 5689 of Lecture Notes in Computer Science, pages 255-356. Springer, 2009.
- [10] Calvanese, D., Keet, C.M., Nutt, W., Rodriguez-Muro, M., Stefanoni, G. Web-based Graphical Querying of Databases through an Ontology: the WONDER System. ACM Symposium on Applied Computing (ACM SAC'10), March 22-26 2010, Sierre, Switzerland.

- [11] Hedman, S. (2004). A first course in logic—an introduction to model theory, proof theory, computability, and complexity. Oxford: Oxford University Press.
- [12] Ian Horrocks, Peter F. Patel-Schneider, and Frank van Harmelen. From SHIQ and RDF to OWL: The making of a web ontology language. *Journal of Web Semantics*, 1(1):7, 2003.
- [13] B. Cuenca Grau, I. Horrocks, B. Motik, B. Parsia, P. Patel-Schneider, and U. Sattler. OWL 2: The next step for OWL. *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, 6(4):309-322, 2008.
- [14] Alessandro Artale, Diego Calvanese, Roman Kontchakov, and Michael Zakharyashev. DL-Lite without the unique name assumption. In *Proc. of the 22nd Int. Workshop on Description Logic (DL 2009)*, volume 477 of CEUR Electronic Workshop Proceedings, <http://ceur-ws.org/>, 2009.
- [15] I. Horrocks, O. Kutz, and U. Sattler. The Even More Irresistible SROIQ. In: *Proc. of the 10th International Conference of Knowledge Representation and Reasoning (KR-2006)*, Lake District UK, 2006
- [16] Franz Baader, Sebastian Brandt, and Carsten Lutz. Pushing the EL Envelope. In *Proc. of the 19th Joint Int. Conf. on Artificial Intelligence (IJCAI 2005)*, 2005.
- [17] Diego Calvanese, Giuseppe de Giacomo, Domenico Lembo, Maurizio Lenzerini, Riccardo Rosati. Tractable Reasoning and Efficient Query Answering in Description Logics: The DL-Lite Family. *J. of Automated Reasoning*, 39(3):385-429, 2007
- [18] There are slides with more details and examples of OWL 2 RL, which are from the ESWC'09 tutorial given by Hitzler, Krotzsch, and Rudolf.
- [19] Benjamin N. Grosz, Ian Horrocks, Raphael Volz, and Stefan Decker. Description Logic Programs: Combining Logic Programs with Description Logic. In *Proc. of the 12th Int. World Wide Web Conference (WWW 2003)*, Budapest, Hungary, 2003. pp. 48-57.
- [20] Keet, C.M. and Rodriguez, M. Toward using biomedical ontologies: trade-offs between ontology languages. *AAAI 2007 Workshop on Semantic e-Science (SeS'07)*, 23 July 2007, Vancouver, Canada. *AAAI 2007 TR WS-07-11*, 65-68.
- [21] Barry Smith. Beyond Concepts, or: Ontology as Reality Representation. Achille Varzi and Laure Vieu (eds.), *Formal Ontology and Information Systems. Proceedings of the Third International Conference (FOIS 2004)*, Amsterdam: IOS Press, 2004, 73-84.
- [22] Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A. WonderWeb Deliverable D18--Ontology library. WonderWeb. 2003. There are also slides of Guarino's OntoLog seminar for a digest of the material.
- [23] Borgo, S., Masolo, C. Foundational choices in DOLCE. In Staab, S., Studer, R. (Eds.), *Handbook on Ontologies (Second Edition)*, Springer Verlag, 2009, p. 361-382.
- [24] Smith B, Ceusters W, Klagges B, Kohler J, Kumar A, Lomax J, Mungall CJ, Neuhaus F, Rector A, Rosse C. Relations in Biomedical Ontologies. *Genome Biology*, 2005, 6:R46.
- [25] Keet, C.M. Constraints for representing transforming entities in bio-ontologies. 11th Congress of the Italian Association for Artificial Intelligence (AI*IA'09), R. Serra and R. Cucchiara (Eds.), Reggio Emilia, Italy, Dec. 9-12, 2009. Springer-Verlag Lecture Notes in Artificial Intelligence LNAI 5883, 11-20.
- [26] Keet, C.M. and Artale, A. Representing and Reasoning over a Taxonomy of Part-Whole Relations. *Applied Ontology*, 2008, 3(1-2): 91-110.
- [27] Artale, A., Guarino, N., and Keet, C.M. Formalising temporal constraints on part-whole relations. 11th International Conference on Principles of Knowledge Representation and Reasoning (KR'08). Gerhard Brewka, Jerome Lang (Eds.) AAAI Press, pp 673-683. Sydney, Australia, September 16-19, 2008
- [28] Varzi, A. Spatial reasoning and ontology: parts, wholes, and locations. In: *Handbook of Spatial Logics*. Berlin Heidelberg: Springer Verlag (2007) 945-1038.

- [29] Randell, D.A., Cui, Z., Cohn, A.G. A spatial logic based on regions and connection. In: Proc. 3rd Int. Conf. on Knowledge Representation and Reasoning, Morgan Kaufmann (1992) 165-176.
- [30] Gruetter, R., Bauer-Messmer, B. Combining OWL with RCC for spatioterminological reasoning on environmental data. In: Third international Workshop OWL: Experiences and Directions (OWLED 2007). 6-7 June 2007, Innsbruck, Austria.
- [31] Presutti, V., Gangemi, A., David, S., de Cea, G. A., Surez-Figueroa, M. C., Montiel-Ponsoda, E., Poveda, M. A library of ontology design patterns: reusable solutions for collaborative design of networked ontologies. NeOn deliverable D2.5.1, Institute of Cognitive Sciences and Technologies (CNR). 2008.
- [32] L. Lubyte, S. Tessaris. Automatic Extraction of Ontologies Wrapping Relational Data Sources. In Proc. of the 20th International Conference on Database and Expert Systems Applications (DEXA 2009).
- [33] Principal issues are described in Hainaut, J.-L., Chandelon, M., Tonneau, C., and Joris, M. (1993). Contribution to a theory of database reverse engineering. In Reverse Engineering, 1993, Proceedings of Working Conference on, pages 161-170.
- [34] Christine Golbreich and Ian Horrocks. The OBO to OWL mapping, GO to OWL 1.1! In Proc. of the Third OWL Experiences and Directions Workshop, number 258 in CEUR (<http://ceur-ws.org/>), 2007. See also the wiki page on `oboInOwl`
- [35] Zhang S, Bodenreider O, Golbreich C. Experience in reasoning with the Foundational Model of Anatomy in OWL-DL. In: Pacific Symposium on Biocomputing 2006, Altman RB, Dunker AK, Hunter L, Murray TA, Klein TE, (Eds.). World Scientific, 2006, 200-211.
- [36] Keet, C.M. Positionalism of relations and its consequences for fact-oriented modelling. International Workshop on Fact-Oriented Modeling (ORM'09), Vilamoura, Portugal, November 4-6, 2009. Meersman, R., Herrero, P., and Dillon T. (Eds.), OTM Workshops, Springer, LNCS 5872, 735-744.
- [37] K. Vila, A. Ferrández. Developing an Ontology for Improving Question Answering in the Agricultural Domain. In: Proceedings of MTSR'09. Springer CCIS 46, 245-256.
- [38] Heiko Dietze, Dimitra Alexopoulou, Michael R. Alvers, Liliana Barrio-Alvers, Bill Andreopoulos, Andreas Doms, Joerg Hakenberg, Jan Moennich, Conrad Plake, Andreas Reischuck, Loic Royer, Thomas Waechter, Matthias Zschunke, and Michael Schroeder. GoPubMed: Exploring PubMed with Ontological Background Knowledge. In Stephen A. Krawetz, editor, Bioinformatics for Systems Biology. Humana Press, 2008.
- [39] Allen H. Renear and Carole L. Palmer. Strategic Reading, Ontologies, and the Future of Scientific Publishing. *Science* 325 (5942), 828. [DOI: 10.1126/science.1157784] (but see also some comments on the paper)
- [40] Dimitra Alexopoulou, Thomas Waechter, Laura Pickersgill, Cecilia Eyre, and Michael Schroeder. Terminologies for text-mining: an experiment in the lipoprotein metabolism domain. *BMC Bioinformatics*, 9(Suppl4):S2, 2008
- [41] Witte, R. Kappler, T. And Baker, C.J.O. Ontology design for biomedical text mining. In: *Semantic Web: revolutionizing knowledge discovery in the life sciences*, Baker, C.J.O., Cheung, H. (eds), Springer: New York, 2007, pp 281-313.
- [42] M. Jarrar, C.M. Keet, and P. Dongilli. Multilingual verbalization of ORM conceptual models and axiomatized ontologies. STARLab Technical Report, Vrije Universiteit Brussels, Belgium. February 2006.
- [43] R. Schwitter, K. Kaljurand, A. Cregan, C. Dolbear, G. Hart. A comparison of three controlled natural languages for OWL 1.1. Proc. of OWLED 2008 DC. Washington, DC, USA metropolitan area, on 1-2 April 2008.
- [44] Keet, C.M. Factors affecting ontology development in ecology. *Data Integration in the Life Sciences 2005 (DILS'05)*, Ludaescher, B, Raschid, L. (eds.). San Diego, USA,

20-22 July 2005. Lecture Notes in Bioinformatics LNBI 3615, Springer Verlag, 2005. pp46-62.

- [45] Dagobert Soergel, Boris Lauser, Anita Liang, Frehiwot Fisseha, Johannes Keizer and Stephen Katz. Reengineering thesauri for new applications: the AGROVOC example. *Journal of Digital Information* 4(4) (2004).
- [46] Maria Angela Biasiotti, Meritxell Fernández-Barrera. Enriching Thesauri with Ontological Information: Eurovoc Thesaurus and DALOS Domain Ontology of Consumer law. *Proceedings of the Third Workshop on Legal Ontologies and Artificial Intelligence Techniques (LOAIT 2009)*. Barcelona, Spain, June 8, 2009.
- [47] Halpin, T., Morgan, T. *Information modeling and relational databases*. 2nd ed. Morgan Kaufmann (2008).
- [48] Keet, C.M. Ontology design parameters for aligning agri-informatics with the Semantic Web. 3rd International Conference on Metadata and Semantics (MTRS'09) -- Special Track on Agriculture, Food & Environment, Oct 1-2 2009 Milan, Italy. F. Sartori, M.A. Sicilia, and N. Manouselis (Eds.), Springer CCIS 46, 239-244.
- [49] Keet, C.M. Dependencies between Ontology Design Parameters. *International Journal of Metadata, Semantics and Ontologies*. (in print).
- [50] Guarino, N. and Welty, C. An Overview of OntoClean. In S. Staab, R. Studer (eds.), *Handbook on Ontologies*, Springer Verlag 2009, pp. 201-220.
- [51] Guarino, N., Welty, C. A formal ontology of properties. In: Dieng, R., Corby, O. (eds.) *EKAU 2000*. LNAI, vol. 1937, pp. 97–112. Springer, Heidelberg (2000)
- [52] Guarino, N., Welty, C. Identity, unity, and individuality: towards a formal toolkit for ontological analysis. In: *Proc. of ECAI 2000*. IOS Press, Amsterdam (2000)
- [53] Parsia, B., Sirin, E., Kalyanpur, A. Debugging OWL ontologies. *World Wide Web Conference (WWW 2005)*. May 10-14, 2005, Chiba, Japan.
- [54] M. Horridge, B. Parsia, and U. Sattler. Laconic and Precise Justifications in OWL. In *Proc. of the 7th International Semantic Web Conference (ISWC 2008)*, Vol. 5318 of LNCS, Springer, 2008.
- [55] Alexander Borgida, Diego Calvanese, and Mariano Rodríguez-Muro. Explanation in the DL-Lite family of description logics. In *Proc. of the 7th Int. Conf. on Ontologies, DataBases, and Applications of Semantics (ODBASE 2008)*, LNCS vol 5332, 1440-1457. Springer, 2008.
- [56] Gomez-Perez, A.; Fernandez-Lopez, M.; Corcho, O. *Ontological Engineering*. Springer Verlag London Ltd. 2004
- [57] Fernandez, M.; Gomez-Perez, A. Pazos, A.; Pazos, J. Building a Chemical Ontology using METHONTOLOGY and the Ontology Design Environment. *IEEE Expert: Special Issue on Uses of Ontologies*, January/February 1999, 37-46.
- [58] Oscar Corcho, Mariano Fernández-López, Asunción Gómez-Pérez, Angel López-Cima. Building legal ontologies with METHONTOLOGY and WebODE. *Law and the Semantic Web 2005*. Springer LNAI 3369, 142-157.
- [59] Corcho, O., Fernandez-Lopez, M. and Gomez-Perez, A. (2003). Methodologies, tools and languages for building ontologies. Where is their meeting point? *Data & Knowledge Engineering*, 46(1): 41-64.
- [60] Mari Carmen Suarez-Figueroa, Guadalupe Aguado de Cea, Carlos Buil, Klaas Dellschaft, Mariano Fernandez-Lopez, Andres Garcia, Asuncion Gomez-Perez, German Herrero, Elena Montiel-Ponsoda, Marta Sabou, Boris Villazon-Terrazas, and Zheng Yufei. NeOn Methodology for Building Contextualized Ontology Networks. *NeOn Deliverable D5.4.1*. 2008.
- [61] Ghidini, B. Kump, S. Lindstaedt, N. Mabhub, V. Pammer, M. Rospocher, and L. Serafini. Moki: The enterprise modelling wiki. In *Proceedings of the 6th Annual European*

Semantic Web Conference (ESWC2009), 2009. Heraklion, Greece, 2009 (demo). LNCS 5554, pp. 831-835.

[62] Jerome Euzenat and Pavel Shvaiko. *Ontology Matching*. Springer. 2007.

[63] Keet, C.M. *Aspects of Ontology Integration*. Technical Report, Napier University, Edinburgh, Scotland. 2004.

[64] Heiner Stuckenschmidt, Christine Parent, Stefano Spaccapietra (Eds.). *Modular Ontologies—Concepts, Theories and Techniques for Knowledge Modularization*. Springer. 2009.

[65] Umberto Straccia. *Managing Uncertainty and Vagueness in Description Logics, Logic Programs and Description Logic Programs*. In *Reasoning Web, 4th International Summer School, 2008*. See also the extensive AAAI'07 tutorial slides by Umberto Straccia.

[66] Thomas Lukasiewicz and Umberto Straccia. 2008. *Managing Uncertainty and Vagueness in Description Logics for the Semantic Web*. *Journal of Web Semantics*, 6:291-308.

[67] Jiang, Y., Wang, J., Tang, S., and Xiao, B. 2009. *Reasoning with rough description logics: An approximate concepts approach*. *Information Sciences*, 179:600-612.

[68] Keet, C.M. *On the feasibility of Description Logic knowledge bases with rough concepts and vague instances*. 23rd International Workshop on Description Logics (DL'10), 4-7 May 2010, Waterloo, Canada.

[69] Alessandro Artale, Christine Parent, and Stefano Spaccapietra. *Evolving objects in temporal information systems*. *Annals of Mathematics and Artificial Intelligence (AMAI)*, 50:5-38, 2007, Springer.

[70] Artale, A., Guarino, N., and Keet, C.M. *Formalising temporal constraints on part-whole relations*. 11th International Conference on Principles of Knowledge Representation and Reasoning (KR'08). Gerhard Brewka, Jerome Lang (Eds.) AAAI Press, pp 673-683. Sydney, Australia, September 16-19, 2008

[71] Alessandro Artale, Roman Kontchakov, Carsten Lutz, Frank Wolter and Michael Zakharyashev. *Temporalising Tractable Description Logics*. Proc. of the 14th International Symposium on Temporal Representation and Reasoning (TIME-07), Alicante, June 2007.

[72] Carsten Lutz, Frank Wolter, and Michael Zakharyashev. *Temporal Description Logics: A Survey*. In: *Proceedings of the Fifteenth International Symposium on Temporal Representation and Reasoning*. IEEE Computer Society Press, 2008.

[73] Frank van Harmelen, Vladimir Lifschitz and Bruce Porter (Eds.). *Handbook of Knowledge Representation*. Elsevier, 2008, 1034p.

[74] Stefan Schulz, Holger Stenzhorn, Martin Boekers and Barry Smith. *Strengths and Limitations of Formal Ontologies in the Biomedical Domain*. *Electronic Journal of Communication, Information and Innovation in Health (Special Issue on Ontologies, Semantic Web and Health)*, 2009.

[75] R. Alberts, D. Calvanese, G. De Giacomo, A. Gerber, M. Horridge, A. Kaplunova, C. Maria Keet, D. Lembo, M. Lenzerini, M. Milicic, R. Moeller, M. Rodriguez-Muro, R. Rosati, U. Sattler, B. Suntisrivaraporn, G. Stefanoni, A.-Y. Turhan, S. Wandelt, M. Wessel. *Analysis of Test Results on Usage Scenarios*. Deliverable TONES-D27, v1.0, Oct. 10 2008.

[76] Keet, C.M., Roos, M. and Marshall, M.S. *A survey of requirements for automated reasoning services for bio-ontologies in OWL*. Third international Workshop OWL: Experiences and Directions (OWLED 2007), 6-7 June 2007, Innsbruck, Austria. CEUR-WS Vol-258.

[77] Carole Goble and Chris Wroe. *The Montagues and the Capulets*. *Comparative and Functional Genomics*, 5(8):623-632, 2004. doi:10.1002/cfg.44