Ontologies: a revamped cross-disciplinary buzzword or a truly promising interdisciplinary research topic?

Peter Spyns and Jan De Bo
Vrije Universiteit Brussel

The following article provides an introductory overview of the different research domains (computational linguistics, terminography, artificial intelligence (AI), philosophy and database semantics) for which ontologies and the emerging field of the Semantic Web have become a main point of interest. It will be pointed out that each of these domains uses a different definition for an ontology. A specific ontology engineering methodology (VUB STAR Lab DOGMA) will be presented and emphasis will be put on the specific role and contribution of (multilingual) terminography in this ontology. In addition, we will explain what ontologies might offer to advance the state of the art of linguistics and terminography.

0. Introduction

The term 'ontology' has almost become a buzzword. After the often cited visionary article by Berners-Lee (2001) on the Semantic Web and its semantic foundation called 'ontology', much research effort and funding have been (and are still being) spent on research topics related to the Semantic Web. Researchers in the domain of knowledge engineering and knowledge management have emerged as the main promoters of ontology research, although ontologies are also highly relevant for other domains such as database semantics, NLP[1], information science (Smith 2003). Consequently, the notion 'ontology' is sometimes interpreted in different ways, resulting in methodological flaws when discussing the various aspects of ontologies. Recently, human language technology research centres (mainly computational linguists and terminologists[2]) have also entered this research area as it has become obvious that their expertise and tools are needed for the successful development of the Semantic Web. In addition to bringing specific knowledge about natural language and terminology to the Semantic Web, research on and tools for ontologies can also be applied by terminologists, linguists and lexicographers alike to support and advance research in their own fields.

In essence, a terminographer's task is to collect all the terms of a technical domain, to provide adequate and unambiguous definitions of their meanings (in many cases, according to the Aristotelian schema of genus and differentiae) and to organise them in semantic networks

(sometimes only in hierarchies). Synonyms can be related to preferred terms (creation of thesauri) in order to stimulate the use of the preferred terms (controlled vocabulary) or additional linguistic characteristics (e.g. gender, part of speech, pronunciation, etc.) can be added to create dictionaries. When relationships (other than hypo-hyperonymy) between terms are given, semantic networks are created. This can be done within one language or across natural languages. Terminographers use mainly all kinds of texts to collect the terms and generally have domain experts validate the definitions.

An ontology engineer's task is to 'organise' a domain conceptually (for some authors this may be limited to single applications), to come up with the relevant concepts and their relationships, to provide explicit and unambiguous definitions of these concepts and to encode these definitions in some formal language so that software applications (e.g. intelligent or autonomous agents) can exchange unambiguous and meaningful messages. Depending on the point of view, model theory (mostly used by database researchers) and proof theory (mostly used by AI researchers) are the main ways of defining the semantics (interpretation function). AI researchers fall back on knowledge elicitation techniques (used to build expert systems) to extract the implicit knowledge from human experts, whereas information system analysts generally use all kinds of reports (flow charts, generated reports, data structure definitions, etc.) to model domains.

One can immediately grasp some of the commonalities and differences between creating terminologies and ontologies. Both disciplines create 'mental organisations' of the domains concerned and both strive for unambiguous communication by providing adequate definitions. The most striking differences may be found in the degree of formality of the vocabulary definitions and the intended outcome. For an ontology engineer, the end result is a commonly agreed upon (formal) set of relevant *definitions* linked to identifying labels (Ushold & King 1995), whereas for a terminologist the collection of domain *terms* with their associated definition is paramount. This distinction is reflected in the audiences which are targeted: software agents versus humans. Nevertheless, one can state that for any domain or situation that necessitates strict definitions of its working terms and notions terminology and ontology can contribute to each other's scientific progress.

The remainder of this paper is organised as follows: in section 1 the various definitions and uses of 'conceptualisation' and 'ontology' are discussed in philosophy (section 1.1), computer science (section 1.2), and linguistics and terminology (section 1.3). These discussions are then followed by elaborations on some pending issues common to all the disciplines mentioned. These issues concern multilinguality, language neutrality, ambiguity and context (section 2). Benefits offered by terminology to ontology (section 3) and vice versa (section 4) precede a section on the VUB STAR Lab DOGMA ontology modelling methodology (section 5).

Related and future work is discussed in section 6. In section **7** we formulate our conclusion.

1. Conceptualisation and Ontology in …

1.1. Philosophy

The term 'ontology' was coined in 1606 by the Swiss philosopher Jacob Lorhard (aka Iacobus Lorhardus)[3]. The title page of his book states *metaphysics, seu ontologia* indicating that 'ontology' refers to "the study of being qua being", as put forward by Aristotle in his book on Metaphysics, IV 1. Ontology as a philosophical subdiscipline concerns the nature and the organisation of reality, and tries to answer questions such as "What are the features common to all beings?" (Guarino & Giaretta 1995: 26). It is the science of what is, of the kinds and structures of objects, properties, events, processes and relations in every area of reality (Smith 2003: 15). The philosophical (and original) definition of the term 'ontology' is still related to the contemporary uses adopted by other sciences.

1.2. Computer science

In accordance with Guarino & Giaretta (1995) and Guarino (1998), we reserve the term *conceptualisation* for the activity of organising or modelling a micro-world, which is reminiscent of the philosophical definition, whereas the term 'ontology' corresponds to the result of encoding the conceptualisation by an ontology language (e.g. RDF(S) (McBride 2004), OWL (Antoniou & van Harmelen 2004)):

> An ontology is a logical theory accounting for the *intended meaning* of a formal vocabulary, i.e. its *ontological commitment*[4] to a particular *conceptualisation* of the world. The intented models of a logical language using such a vocabulary are constrained by its ontological commitment. An ontology indirectly reflects this commitment (and the underlying conceptualisation) by approximating these intended models. (Guarino 1998: 7) [his emphasis]

Smith (2003: 162) points to the reductionist nature of a model as defined by Gruber (1995), i.e. only characteristics and entities of a domain relevant for a certain purpose are modelled (the adequacy relationship with the entire actual reality is no longer of concern). Nevertheless, conceptualisations (and hence ontologies) gain in quality and consistency if a truthful relationship with reality is maintained. This, in turn, will help to ensure the unifiability of separate ontologies (Smith 2003: 163).

1.2.1. Information systems

One of the information analyst's tasks is to model a system and to describe functions that are needed by companies (or organisations) to achieve (business) goals. A model should capture the essential information (characteristics, entities, roles) of the organisation (the information model or infological schema). Information analysts eventually produce conceptual models that form the bases for database schemas. The interoperability problem manifests itself when, for example, two separately developed database applications have to exchange information while both using different vocabularies. In the realm of application A, an entity is represented using label X, while label Y refers to the same entity in the micro-world of application B. This is sometimes dubbed "the tower of Babel problem" (Smith 2003: 158). Recently, electronic business exchanges and transactions between computers have been realised using Web Services, which exacerbates the problems because of idiosyncratic (business) vocabulary.

1.2.2. Artificial Intelligence (AI)

Knowledge engineering in AI has been involved very much with representing facts (static aspects, declarative knowledge) and rules (dynamic aspects, procedural knowledge) that make up and govern a micro-world. Reasoning and inferring new knowledge (and triggering actions) from the set of rules and the current state of the micro-world are at the heart of this kind of systems. Contrary to the early solipsist AI systems (closed world assumption) the Semantic Web is fundamentally an open system, so here again interoperability of autonomous systems (intelligent software agents) has become a prerequisite. An ontology is believed to offer the answer to these interoperability problems.

1.3. Linguistics and terminology

From the Ancient Greeks until now the relation between language and the ways in which people organise their worlds (or *the* world) have been studied intensively. This relation was crystallised in the Sapir-Whorf hypothesis (linguistic determinism versus linguistic relativity). The hypothesis stated that a language and the organisation of a world (be it *the* world, an individual's world or a system's micro-world) are related in a strong (deterministic) or weaker (relative) way. Defining a linguistic theory or giving a terminological account of a domain is also a conceptualisation. Formalising and implementing a linguistic theory has become widespread nowadays, but representing the formal definitions by means of an ontology is quite rare. In the same vein, lexicologists and termino-

graphers formally defining and implementing their frameworks (the meta-languages of their dictionaries or terminological systems) is still a rather recent phenomenon (Farrar et al. 2002; Lenci et al. 2002; Vouros & Eumeridou 2002).

2. Multilinguality, language neutrality, ambiguity and context

Although we have not mentioned it explicitly, *multilinguality* constitutes an important issue. From a philosophical point of view, it supports the principle of linguistic relativity (i.e. speaking a different language may result in thinking differently about the same world). For linguists and terminologists, in opposition to ontology engineers, multilinguality is relatively well understood. SIMPLE (Vouros & Eumeridou 2002), for example, is an attempt to create an upper ontology for the linguistic domain based on the grammatical framework of twelve natural languages.

Some (language) philosophers advocated the notion of *language neutral* concepts (Hovy & Nirenburg 1992) instead of stating that concepts are language-independent. The basic idea is that, even if language-independent concepts are desirable/favoured, in practice (natural) language biases will inevitably slip in through the modeller's language. Therefore, the biases introduced should be neutral (i.e. they should not introduce semantic distinctions based on idiosyncratic distinctions) to as many languages as possible. Hovy and Nirenburg (1992: 5) note that language neutrality can only be approached asymptotically. They propose a stepwise folding in of one language at a time. Some authors claim that for technical domains, language neutrality could be achieved (Hirst 2004: 225). Guarino (1998: 4) states that ontologies can differ in vocabulary (e.g. using English or Italian words) while sharing the same conceptualisation[5]. However, this citation also reveals a misconception often made by ontology researchers: they use a natural language (NL) word as a concept label for convenience but at the same time forget that in doing so the distinction between the language and the conceptual levels is blurred. In addition, *term* and *vocabulary* are used to speak about the logical terms and vocabulary of a (first order) ontology language (i.e. its signature) and not about the technical natural language terms and vocabulary. So it is somewhat absurd to read about translating ontologies with the Altavista Babelfish tool (Sure 2003: 105) as it involves more than simply translating NL terms. This also means that results of text or web mining cannot simply be used as direct input to an ontology. One could argue that if a terminologist has done a good job, the NL technical terms lexicalise the underlying concepts with a one-to-one mapping. Unfortunately, this is an ideal case, and the examples of ontologies found in the literature contain many commonly used non-technical terms. Consequently, the shared and common agreement typical of an ontology relies largely on the way in which humans intuitively understand the NL terms. On this issue we fully agree (see De Bo et al. 2003) with

Nirenburg and Raskin that some scholars persist in this natural-language fallacy positively, as it were, by insisting on using natural-language words instead of ontological concepts to represent natural language meanings (2001: 153).

    These latter authors also point out that the difference between an ontology (or conceptualisation) versus a natural language lies not so much in the presence or absence of ambiguity, but rather in the defined and consensual nature of the ontological concepts and their labels (Nirenburg & Raskin 2001: 157). Consequently, an ontology may have *ambiguous* terms. But as a computer has no means of detecting the ambiguity, they argue that there is no point in worrying about potential ambiguities. We would like to proceed more cautiously. As ultimately a human will be in the loop (even via the intermediary of computers and intelligent agents), we consider it more prudent to devise an ontology engineering methodology that avoids ambiguities as much as possible, especially with an eye on ontology integration, which gives rise to a range of potential mismatch types (Tamma 2001). The notion of context plays an important role in this issue.

3. Linguistics and terminology for ontologies

Using the information above a terminographer's main task can be described as collecting the technical vocabulary of a domain, preliminarily organising the technical terms (e.g. in terms of synonymy, causality (Cabré et al. 2004)) and providing clear definitions for these terms. The terminographer will use various tools (in particular mining tools, which automate term collection and synonym grouping (Gargouri et al. 2004)) to speed up the work and which support validation of the outcomes. Inevitably, a set of seed terms (corresponding to an upper ontology or to the top of a domain ontology) will be needed to start the process. Seed terms can be given by domain experts or can be discovered by applying statistical techniques and measures (Gillam & Tariq 2004) or unsupervised mining methods (Reinberger et al. 2004). Terminographers are also trained in locating, gathering and preprocessing the necessary information sources for specific technical domains.

    Providing good definitions for the (preferred) terms that the various stakeholders in the domain can agree upon is the key activity. In addition, when technical terms from different languages are to be compared and combined, the terminographer has to take care in foreseeing how the words from the various languages relate to one another. Some authors also propose taking cultural differences into account (Temmerman 2003). The terminologist should keep track of the specific documents (or maybe even sentences) in which particular technical terms are used (or some pointers to the occurrences of the technical terms that are deemed representative). It is thus rather obvious that ontology engi-

neers can learn from terminologists and linguists, even if only to avoid modelling pitfalls which terminologists and linguists have already encountered on their scientific journey spanning more than two millennia.

4. Ontologies for linguistics and terminology

Research on ontologies brings a new meta-level framework into the work of terminographers, lexicologists and linguists. The object of the research are the various theoretico-linguistic frameworks in the sense that the descriptive notions (e.g. singular, plural, dual, etc.) are catalogued, defined and organised in models (or semantic networks). Initiatives of this kind have already started – e.g. EAGLES (Sanfilippo 1998). The additional step is to formally describe and constrain these notions. A noun, for example, *may* express case marking and always expresses number (optionality). The antonym relation *necessarily* involves *two* notions (cardinality and mandatoriness). This model can be used as a reference to store linguistic resources in databases, and subsequently as a basis for concept-based retrieval, including concept-driven query interfaces. For instance, one could query the dictionary for entries that (transitively) are part of the same whole (this query requires inferencing). Already existing linguistic resources can map their internal meta-level linguistic vocabulary to the definitions of the ontology (= ontologically committing to the conceptualisation). To our knowledge, there are no large ontology-driven linguistic resources – as described above – available yet.

There will be no single overall linguistic ontology[6], as there is no single world ontology, as there is no single business ontology, etc. Machine readable dictionaries, thesauri, term banks, etc. stored in databases will select a specific ontology and conform to it. Consequently, standard information system interoperability problems between various databases and applications (here containing various linguistic resources) can be overcome. Ontologies isolate semantics from applications (meaning independence) in the same way in which databases isolate data structures from applications (data independence) (Meersman 2001b). In addition, linguists will be able to re-use the various formalisms, tools, software components and architecture that have been and are being created in the context of the Semantic Web. One can easily imagine a portal, including semantic web services, which offers multilingual terminological information. In this respect, terminographers profit from the scientific progress in database, AI and internet technologies.

5. DOGMA: Developing Ontology Guided Mediation for Agents

A DOGMA-inspired ontology is defined in a logical sense, i.e. as a "representationless" mathematical object which forms the range of a classical

interpretation mapping from a first order language (assumed to represent an application lexically) to a set of possible ('plausible') conceptualisations of the real world domain. This definition also leads to methodological approaches that naturally extend database modelling theory (Meersman 2001a) and practice (Meersman 2001b). We have introduced the double articulation of an ontology (Spyns et al. 2002) by decomposing it formally into an *ontology base* and into instances of their explicit ontological commitments. The latter become reified in our architecture as a separate mediating layer called *commitment layer*. Recently (De Bo et al. 2003) the DOGMA framework has been refined to explicitly add the distinction between the language and conceptual levels by formalising the context and introducing language identifiers.

- An ontology base consists of intuitively plausible conceptualisations of a real world domain, i.e. specific binary fact types called *meta-lexons* and formally noted as <$concept_1$ – relationship – $concept_2$>. They are derived from *lexons*, written as <(γ,λ): $term_1$, role, co-role, $term_2$>. Informally we say that a *lexon* is a fact that may hold for some application, expressing in that case that within the context γ and for the language λ the $term_1$ may plausibly have $term_2$ occur in *role* with it (and inversely $term_2$ maintains a *co-role* relation with $term_1$). Lexons are independent of specific applications and should cover relatively broad domains[7] (linguistic level). Lexons constitute a lexon base, which is constituted by lexons grouped by context and language. Meta-lexons are language-independent and context-independent (conceptual level). Terms are mapped to concepts (word senses) via the context-language combination – see De Bo et al. (2003) for more details)[8]. The same process applies to a (co-)role and a relationship.

- The *layer of ontological commitments* mediates between the ontology base and its applications. The commitment layer is organised as a set of ontological commitments, each being an explicit instance of an (intensional) first-order interpretation of a task in terms of the ontology base (Jarrar & Meersman 2002). Each commitment is a consistent set of rules (or axioms) in a given syntax which adds specific semantics to a selection of meta-lexons of the ontology base. Commitments have a varying degree of genericity. Sets of ontological commitments can be regarded as re-usable knowledge components.

Based on the reasoning above, we propose to have the ontology engineering process done in two major steps: (i) a linguistic step and (ii) a conceptual step. In this paper, we will not detail the creation of a commitment, i.e. adding formal semantic constraints on (parts of) the language-independent domain conceptualisation. DOM[2] heavily relies on Object Role Modelling (Halpin 2001), a conceptual database modelling methodology, which we have adapted for ontology modelling. However,

DOM[2] still lacks aspects of distributed collaborative modelling. We hope that we can draw upon existing practices from the terminology community to refine the method.

- The *linguistic* step is to be taken in close contact with the domain specialists and terminologists. The domain specialists will provide the ontology modeller with the necessary documents, useful insights, domain expertise and so on. The terminologists help the domain specialists in selecting terms and creating lexons. The knowledge needed for these steps is 'carried' by a language (of the documentation, of the domain specialists, as already used in database schemas or programs, etc.).

- The *conceptual* step is a combined job for the terminologist and ontology engineer. The former is responsible for finding (or creating) the most adequate definitions and concept (and relationship) labels for terms, and the latter for organising the domain model using these concepts and relationships. During this step, one passes from the linguistic level to the conceptual level: a lexon becomes a meta-lexon. Of course, the results have to be validated by domain experts and/or other stakeholders' representatives.

Various tools can be of assistance to collect terms, to produce lexons or for preparatory activities. MindManager™, for example, is used in the On-to-Knowledge methodology (Sure 2003) to make preliminary domain conceptualisations. These tools are specifically helpful during initial brainstorming or exploratory sessions. Various other sources (e.g. DB schemas, XML DTDs, organisation charts, etc.) can be used to mine or build ontologies.

DOM[2] step 1: verbalising information examples as elementary facts
In many cases, we start from scratch with a specific application at hand. As, by definition, ontologies should be shared, care should be taken not to limit the domain world to the one of the application. On the other hand, one must beware of modelling the 'entire world'. By preference, the data sources collected for the domain ontology should already have an agreed and common character (e.g. standard text, reference classification, etc.). If these sources do not exist, domain specialists and other stakeholders have to agree whether or not to include specific entity types.
    The first step is to begin with familiar examples of relevant information, as can be found in textual descriptions, and to express these examples as elementary facts. The complexity of textual information must be reduced to simple sentences expressing elementary facts. An elementary fact is a simple assertion, or atomic proposition, about the universe of discourse. They are simple assertions stating that particular objects play particular roles. As mentioned by Halpin (2001: 61) an ele-

mentary fact cannot be split into smaller units of information. As long as a sentence contains words such as 'and', 'or', 'if', 'not', 'all' or 'some', it does not express an elementary fact.

DOM² step 2: creating lexons (per context and language)
One hopes and expects that NLP techniques will be able to deliver lexons after text processing, just as reverse database engineering techniques could do for existing databases and conceptual models. In the absence of automated techniques, it is better to choose existing natural language words (or combinations) for terms and roles of a lexon. Many modelling approaches express roles or relationships using verbs. During this stage, domain experts organise their worlds intuitively and informally. Nevertheless, the less ambiguity there is about the intuitive meanings of words used for the terms and roles, the better. By carefully choosing a role name, for example, a modeller can indicate that the role might have a transitive nature.

DOM² step 3: creating meta-lexons
Meta-lexons are created by 'replacing' the language terms or words (i.e. the terms and roles for a specific language and context) of the lexons with labels identifying concepts and conceptual relationships. Concept definitions might already be available (e.g. word sense descriptions in WordNet (Fellbaum 1998), technical terms and definition collections) and thus commonly agreed upon and/or generally accepted. Others have to be construed on the spot as WordNet mostly covers non-technical vocabulary. We recommend doing this in the same format and style as WordNet, thereby also creating (multilingual) synsets. Domain experts, ontology engineers and terminologists have to work together to achieve this. As concepts and relationships stand for a unique notion or sense, the context and language identifiers become superfluous – see Spyns et al. (2004) for more details.

6. Related and future work

Not so many ontology modelling methodologies are mentioned in the literature – see Gómez-Pérez (2004: 157-163) for an overview. Hardly any attention is paid to the difference(s) between the linguistic level and the conceptual level. At the same time, not many linguists are developing ontologies. We have already mentioned them in the previous sections. To our knowledge, we have encountered only one example of a methodology (called termontography) that explicitly seeks to combine terminological modelling with ontological modelling (Temmerman & Kerremans 2003). However, we disagree with Temmerman and Kerremans in what they call the categorisation framework. To the extent that an existing conceptualisation can be reused as categorisation framework, we can follow their

approach. In the opposite direction, creating the categorisation framework that exhibits the level of detail as shown in their example – in opposition to upper ontologies – is already creating the conceptualisation and associating definitions. This leads to a Catch-22 situation: terms are needed to create the concepts (what Temmerman and Kerremans call units of understanding), but then concepts are needed to link the terms to.

We propose the use of existing (technical) term banks or resources such as WordNet to link terms (potentially collected automatically) to definitions or word senses – existing ones if available, newly created ones if needed. It should not be forgotten that this is meant for human users of the Semantic Web, and these users are not necessarily the Semantic Web's primary targeted users (they are, however, the most determining ones). The agreement on the definition is more important than the choice of a particular term.

In the future, we hope to be able to apply principles from the discipline of formal ontology engineering (Guarino & Welty 2004), as well as from the discipline of terminography (especially the collaborative and modelling preparatory aspects). The aim is to come up with a cookbook-style manual for ontological engineering. Using it to build an ontology for terminographical applications is an interesting test case. This will partly be done during the OMTFI project, which is a collaboration between the Vrije Universiteit Brussel – STAR Lab and the Erasmushogeschool Brussel – CVC. Also, the EU Leonardo da Vinci project called "Co-Drive" (BE/04/B/F/PP-144.339) has a work package to implement tools relating multilingual terms with appropriate word senses.

7. Conclusion

In this paper, we have given an overview of various definitions and uses of ontologies to allow the reader to distinguish genuine usage of the term from buzzword usage. In addition, we have presented how ontology and terminology (or by extension linguistic engineering) can benefit from scientific progress and insights in each discipline. This has been illustrated by discussing some specific issues such as ambiguity, language independence, and multilinguality. Subsequently, a particular ontology modelling methodology has been partly sketched. It explicitly distinguishes a linguistic level from a conceptual level. A multidisciplinary collaboration between linguists, computer scientists and domain experts is needed. As few related methodologies exist, a potentially promising cross-disciplinary research domain awaits further exploitation.

Bibliography

Antoniou, G. & F. van Harmelen (2004). "Web Ontology Language: OWL". S. Staab and R. Studer (2004), 67 - 92.

Berners-Lee, T., J. Hendler & O. Lassila (2001). "The Semantic Web. A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities". *Scientific American*, May 2001.

Cabré, M.-T., J. Feliu & J. Vivaldi (2004). "Base de connaissances GENOMA: le rôle de l'ontologie". F. Ibekwe-San Juan and S. Lainé-Cruzel (2004).

De Bo, J., P. Spyns & R. Meersman (2003). "Creating a "DOGMAtic" multilingual ontology infrastructure to support a semantic portal". R. Meersman, Z. Tari et al. (eds) (2003). *On the Move to Meaningful Internet Systems 2003: OTM 2003 Workshops*, LNCS 2889, Heidelberg: Springer Verlag, 253 – 266.

Farrar, S., W.D. Lewis & D.T. Langendoen (2002). "A common ontology for Linguistic Concepts", *Proceedings of the Knowledge Technology Conference* http://emeld.org/documents/KnowTech-CommonOntology.pdf (consulted 11/08/2004).

Fellbaum, C. (ed.) (1998). *Wordnet, An Electronic Lexical Database,* Cambridge: The MIT Press.

Gargouri, Y., B. Lefebvre & J.-G. Meunier (2004). "ONTOLOGICO: vers un outil d'assistance au développement itératif des ontologies". F. Ibekwe-San Juan & S. Lainé-Cruzel.(2004)

Gillam, L. & M. Tariq (2004). "Ontology via Terminology?" F. Ibekwe-San Juan & Lainé-Cruzel (2004).

Gómez-Pérez, A, M. Fernández-López & O. Corcho (2004). *Ontological Engineering with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web.* Heidelberg: Springer Verlag.

Gruber, T. (1995). "Towards Principles for the Design of Ontologies Used for Knowledge Sharing" *International Journal of Human-Computer studies* 43(5/6), 907 – 928.

Guarino, N. & P. Giaretta (1995). "Ontologies and Knowledge Bases: Towards a Terminological Clarification". N. Mars (ed.) (1995). *Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing,* Amsterdam: IOS Press, 25 – 32.

Guarino, N. (1998). "Formal Ontologies and Information Systems". N. Guarino (ed.) (1998). *Proceedings of FOIS98*, Amsterdam: IOS Press, 3 – 15.

Guarino, N. & C. Welty. (2004). "An Overview of OntoClean". S. Staab & R. Studer (2004), 151 – 172.

Halpin, T. (2001). *Information Modeling and Relational Databases: from conceptual analysis to logical design*. San Francisco: Morgan-Kaufmann.

Hirst, G. (2004). "Ontology and the Lexicon". S. Staab & R. Studer (2004), 209 – 229.

Hovy, E. & S. Nirenburg (1992). "Approximating an interlingua in a principled way". *Proceedings of the DARPA Speech and Natural Language Workshop*, http://www.isi.edu/natural-language/people/hovy/papers/92darpa-il.pdf (consulted 11/08/2004).

Ibekwe-San Juan F. & S. Lainé-Cruzel (eds) (2004). *Proceedings of the Workshop on Terminology, Ontology and Knowledge Representation* http://www.univ-lyon3.fr/partagedessavoirs/termino2004/programgb.htm (consulted 11/11/2004).

Jarrar, M. & R. Meersman (2002). "Formal Ontology Engineering in the DOGMA Approach". R. Meersman, & Z. Tari et al. (eds) (2002). *On the Move to Meaningful Internet Systems 2002: CoopIS, DOA, and ODBASE*, LNCS 2519, Heidelberg: Springer Verlag, 1238 – 1254.

Kerremans, K. (2004) "Categorisation frameworks in termontography". *Linguistica Antverpiensia* 3/23004.

Lenci, A., N. Calzolari & A. Zampolli (2002). "From Text to Content: Computational Lexicons and the Semantic Web". *AAAI Workshop on Semantic Web meets language resources* [related to the 18th National Conference on Artificial Intelligence].

McBride, B. (2004). "The Resource Description Framework (RDF) and its Vocabulary Description Language RDFS". S. Staab & R. Studer (2004), 51 – 66.

Meersman, R. (2001a). *Reusing certain database design principles, methods and techniques for ontology theory, construction and methodology*, STAR Lab Technical Report STAR-2001-01, Brussel
http://www.starlab.vub.ac.be/publications/STAR-2001-01.pdf (consulted 11/11/2004).

Meersman, R. (2001b). "Ontologies and Databases: More than a Fleeting Resemblance". A. d'Atri & M. Missikoff (eds) (2001). *OES/SEO 2001 Rome Workshop*, Luiss Publications.

Nirenburg, S. & V. Raskin (2001). "Ontological Semantics, Formal Ontology, and Ambiguity". *Proceedings of the Second International Conference on Formal Ontology in Information Systems*. ACM Press, 151 – 161.

Reinberger, M.-L., P. Spyns,A.J. Pretorius & W. Daelemans (2004). "Automatic initiation of an ontology". R. Meersman, & Z. Tari et al. (eds), *On the Move to Meaningful Internet Systems 2004: CoopIS, DOA, and ODBASE (part 1)*, LNCS 3290, Heidelberg: Springer Verlag, 600 – 618.

Smith, B. (2003). "Ontology", L. Floridi (ed.) (2003). *Blackwell Guide to the Philosophy of Computing and Information.* Oxford: Blackwell, 155 – 166 (see also http://ontology.buffalo.edu/smith/articles/ontologies.htm for an extended version) (consulted 11/08/2004).

Sanfilippo, A. et al. (1998). "EAGLES Preliminary Recommendations on Semantic Encoding". The EAGLES Lexicon Interest Group.

Sowa, J. F. (2000). *Knowledge Representation: Logical, Philosophical, and Computational Foundations*, Brooks Cole Publishing Co.

Spyns, P., R. Meersman & M. Jarrar (2002). "Data modelling versus Ontological engineering". *SIGMOD Record* 31(4), 7 – 12.

Spyns, P., S. Van Acker & M. Wynants (2004). "Using a novel ORM-based ontology modelling method to build an experimental Innovation router". E. Motta, N. Shadbolt, A. Stutt & J. Gibbins (eds) (2004). *Engineering Knowledge in the Age of the Semantic Web*, LNAI 3257, Heidelberg: Springer Verlag, 82 – 98.

Sure, Y. (2003). *Methodology, Tools & Case Studies for Ontology based Knowledge Management*. PhD thesis. Institute AIFB at Karlsruhe.

Staab, S. & R. Studer, (eds) (2004). *Handbook on Ontologies*, Heidelberg: Springer Verlag.

Tamma, V. (2001). *An Ontology Model supporting Multiple Ontologies for Knowledge sharing*. PhD thesis. University of Liverpool. [http://www.csc.liv.ac.uk/~valli/Thesis.zip] (consulted 08/2004).

Temmerman, R & K. Kerremans (2003). "Termontography: Ontology Building and the Sociocognitive Approach to Terminology Description". *Proceedings of the CIL-17 Conference.*

Ushold M. & M. King (1995). "Towards a Methodology for Building Ontologies". *Proceedings of the IJCAI95 Workshop on Basic Ontological Issues in Knowledge Sharing.*

Vouros, G & E. Eumeridou (2002). "Simple and EuroWordNet: towards the Prometheus ontological framework". *Terminology* 8(2), 245 – 284.

---

[1]   In some cases (e.g. NLP) research into ontologies was already taking place even if the term 'ontology' was not being used yet.

[2]   The words 'linguist', 'lexicologist' and 'terminologist' will be considered as synonyms for the sake of simplicity and will thus be used interchangeably.

[3]   Discovered by Raul Corazzon (http://www.formalontology.it/history.htm (consulted 11/08/2004)). Until then, Rudolf Göckel (aka Rudolfus Goclenius) was cited as the first to have coined the term *ontology* (in 1613).

[4]   Smith (2003: 166) defines the ontological commitment of a theory (or of an individual or a culture) as consisting in the objects or types of objects which the theory (or individual or culture) assumes to be in existence.

[5]   This suggests a universalist view (opposed to the Sapir-Whorf hypothesis): the same thought can be expressed in various ways in various languages.

[6]   Not to be confused with WordNet (Fellbaum 1998), which is also called a linguistic ontology by some authors.

[7]   In our opinion, a purpose-independent or task-independent ontology can never be created (due to, for example, the granularity applied – also see Sowa (2000: 171)). Thanks to a collaborative modelling process, however, an ontology should attain a high degree of application independence while somehow remaining limited to a 'family' of applications.

[8]   Other factors (e.g. the socio-linguistic background) could be added to refine the mapping relationship. Kerremans (2004) proposes the cultural environment.