Ontology Engineering: A Reality Check

Elena Paslaru Bontas Simperl[1] and Christoph Tempich[2]

[1]Free University of Berlin, Takustr. 9, 14195 Berlin, Germany
paslaru@inf.fu-berlin.de
[2]Institute AIFB, University of Karlsruhe, 76128 Karlsruhe, Germany
tempich@aifb.uni-karlsruhe.de

Abstract. The theoretical results achieved in the ontology engineering field in the last fifteen years are of incontestable value for the prospected large scale take-up of semantic technologies. Their range of application in real-world projects is, however, so far comparatively limited, despite the growing number of ontologies online available. This restricted impact was confirmed in a three month empirical study, in which we examined over 34 contemporary ontology development projects from a process- and costs-oriented perspective. In this paper we give an account of the results of this study. We conclude that ontology engineering research should strive for a unified, lightweight and component-based methodological framework, principally targeted at domain experts, in addition to consolidating the existing approaches.

1 Introduction

The emergence of the Semantic Web has marked an important step in the evolution of ontologies. Regarded as a means for a shared knowledge understanding and a way to (formally) represent real world domains, they are expected to play a crucial role in data and application integration at public and corporate level. In the last decades researchers have proposed process methodologies for various ontology engineering scenarios [7]. Given the difficulties related to building and maintaining ontologies, a methodological framework provides important benefits: it structures the process, thus breaking its complexity down to manageable tasks, clarifies the responsibilities of the process participants, increases its traceability and enables systematic quality assurance procedures.

The theoretical results achieved in the ontology engineering field in the last fifteen years are of incontestable value for the prospected large scale take-up of semantic technologies. Their range of application in real-world projects is, however, so far comparatively limited, despite the growing number of ontologies online available.[1] This restricted impact was confirmed in a three month empirical study, in which we surveyed 34 recent ontology engineering projects from industry and academia in order to give an account of the current ontology engineering practice and of the efforts involved in these activities. The study focused on process-related rather than modeling issues; in particular it analyzed

---

[1] Refer for example to `http://swoogle.umbc.edu/` for recent statistics on this topic.

the impact of actual research achievements on real world ontology engineering projects, the complexity of particular ontology building tasks, the quality of the support tools, and the various usage scenarios for ontologies.

The majority of the investigated case studies did not follow a systematic approach to ontology building, because the participants each underestimated the associated efforts or were not aware of the availability of methodological support. Nevertheless, the inventory of activities carried out in each of the ontology engineering projects largely overlapped with the ones described in the literature—though the concrete order of execution of these activities or the way they were combined were not necessarily the same as foreseen within academic process descriptions.

Accounting for the experiences gained during the survey, this paper argues that there is a need for a unified ontology engineering process model—following the analogue development in the software field, which moved from a multitude of different process models towards the idea of *method engineering*. Complementarily to initiatives aiming to outreach existing results to the industry, ontology engineering research should consider aligning previous disparate efforts in order to provide real added value to the community of domain experts building ontologies. Methodologies should offer their applicants support at a level of detail which is adjusted to the complexity of each ontology development activity and to the challenges of the project setting. They should be customizable to various special needs, such as, the learning of ontologies from text, and should concentrate on providing a comprehensive range of methods which can be arbitrarily combined and exchanged rather than postulating static process models.

The remainder of this paper begins with a review of analytical and empirical evaluations of existing methodologies in Section 2. After a brief description of ontology engineering processes in Section 3, we present our survey and discuss its results in Section 4. We draw conclusions for future research in Section 5. Section 6 summarizes our work and concludes this paper.

2 Related Work

This section introduces surveys on ontology engineering methodologies and empirical ontology engineering case study descriptions previously published in the literature.

2.1 Analytical Surveys

This paper is concerned with the impact of existing methodologies on current ontology engineering practice. Previous work primarily focused on the analytical evaluation of these approaches. The surveys defined a number of criteria derived from various sources and evaluated methodologies accordingly. By contrast, our considerations are based on case study experiences distilled from multiple expert interviews.

[11] summarizes the main ontology engineering activities covered by the methodologies available at that time. The authors identify the need for guidance on ontology reuse.[2] They further demand that ontology engineering methodologies should not be based on singular project experiences, but should be applied in more settings in order to claim generality. They argue that methodologies become useful for practitioners only with a larger record of projects successfully carried out. They conclude that most methodologies offer some guidance on the major engineering activities, but that there is plenty of scope for refinement. [3] compare different ontology engineering methodologies w.r.t. the granularity of their process descriptions. They introduce a series of ontology engineering activities, classified in the categories "ontology management", "ontology development" and "ontology support", and analyze which methodologies implement which activities to which extent. They conclude that no methodology covers all required ontology engineering activities, and that tool support is still missing for most of the analyzed methodologies (due to 2003). More recently, [25] identified requirements on ontology engineering methodologies to support the development of ontologies for knowledge management applications. The authors compare the support offered by existing methodologies against these requirements and outline a number of open issues for further research and development. In particular, they emphasize that current methodologies are not integrated into classical business process models and do not take into account this dimension to a satisfactory extent.

In summary, analytical surveys identify open issues for ontology engineering methodologies *from a theoretical perspective*. This is orthogonal to our empirical approach, which examines the utility of existing methodologies for current practice, thus forming the basis for new analytical evaluations in the future.

2.2 Empirical Studies

In the following we give an overview of the most prominent case studies related to ontologies which have been published in the Knowledge/Ontology Engineering literature from the early nineties to now. Claiming by no means for completeness, this overview concentrates on *empirical* studies reporting on concrete experiences in developing or deploying ontologies—with or without the help of a specific methodological framework. Our aim is to point out the practical conclusions, lessons learned and guidelines derived from these studies, in order to endorse and complete the results of our own investigations.

The case studies can be classified according to two dimensions: the method employed to construct the ontology, and the purpose of the experiment. According to the former we can distinguish among those aiming at building ontologies i) from scratch, ii) by reuse or iii) with the help of (automatic) knowledge acquisition techniques. The objectives followed by the studies are twofold: the majority of the experiments have been carried out to validate a particular methodology

---

[2] By 1998 no methodologies for ontology reuse had been proposed yet.

or method, or to exemplify the usage of a specific tool for ontology engineering; a considerably lower number of studies applied existing results in ontology engineering as methodological or technological support for creating a specific ontology. This last category of studies gives an account of the impact, the usability and the added value of current ontology engineering research and development in real-world settings.

[2, 5, 6, 14, 15, 27, 29], to name only a few, report on the application of self-developed methodologies and methods to manually build different types of ontologies. The results of these experiments are centered on the (positive) usability of the proposed approach in the designated context and marginally address the question of process operationalization. With this respect each paper emphasizes the need for high quality tools (e.g. for translating, matching or merging, to name the most frequently mentioned ones). Furthermore the authors acknowledge the resource-intensive nature of the manual application of the proposed approaches, and propose (or highlight the need for) dedicated ontology engineering environments. More details on this type of empirical studies can be found for example in [7].

In the remaining of this section we overview several case studies *applying* existing ontology engineering research and development to real-world settings. Some of these experiments are situated in the originating context of particular methodologies and methods, therefore resorting to these as guidance for the completion of the ontology construction tasks (e.g., the ontologies introduced in [1, 13, 21, 24, 30]). The experiments primarily consist of a description of the engineering process followed by superficial observations related to the lack of adequate technological support, at most. By contrast, other experiments evaluate several methodologies and methods w.r.t. their relevance and usability, prior to applying them in a particular application setting, or operate the engineering process without nominally committing to existing techniques [10, 12, 17, 18, 22, 26, 28]. The results of these evaluation procedures reveal the limited usability or the poor impact of the most part of existing ontology engineering methodologies and methods.

Uschold and Healy report on an experiment in which an engineering mathematics ontology is used to detail the specification of a simple software tool and to allow units-conversion and dimensional consistency checking capabilities to this application [28]. In this attempt they tackle some of the most important issues related to ontology engineering processes, from ontology evaluation to more technical activities such as the translation to new representation languages and the integration of multiple ontologies to a new application setting. The case study reported in [28], though not investigating the complete ontology life cycle, reveals several important limitations of ontology-driven research: the difficulties of automatic translation between representation formats and the need for scalable and efficient technologies. Russ and colleagues describe a case study in which an ontology covering the air campaign domain was built by reusing existing ontologies partially covering its context [22]. The conclusions of this case study are comparable to the ones stated by Uschold and Healy: while reusing ontologies was perceived to be beneficial for this particular setting, the authors emphasize the limitations of the techniques available so far, in particular related to language translators and ontology merging. The case study in [17] reports on the feasibility of current technologies in managing and using large scale medical ontologies, emphasizing the need for a more task-oriented approach to ontology engineering at methodological level, and the lack of feasible methods for extracting ontological knowledge from semi-structured models. Paslaru and Mochol point out the limitations of current technologies for translating, comparing and merging ontologies in [18], as resulted from a case study in which a Human Resources ontology was built on the basis of standard eRecruitment classifications. Challenges related to organizational aspects of ontology engineering processes are mentioned in [10, 12, 26].

Summarizing existing methodologies were applied in a small number of in situ case studies under the direct supervision and with the participation of the methodology developers, at most. The methodologies were only used in the application scenarios they were originally designed for, and little is known about their use for related application areas.

3 Ontology Engineering in a Nutshell

Ontology Engineering (OE) is formally defined as "the set of activities that concern the ontology development process, the ontology life cycle, and the methodologies, tools and languages for building ontologies"[7]. This section summarizes the most important of these activities.

Ontology engineering methodologies support ontology building for centralized ontology applications.[3] [4, 20] focus on the consensus building process in collaborative ontology engineering. Methodologies guiding the ontology reuse process, *e.g.*, [6, 19] or the ontology learning process, *e.g.*, [16] complete the picture.

Methodologies divide the ontology building process in a varying number of stages, and propose a number of activities for each stage. The importance of a particular activity within a methodology primarily depends on, *e.g.*, the characteristics of the ontology-based application, the complexity of the ontology to be built, the availability of information sources, and the experience of the ontology engineers.

[7] differentiates among *management*, *development-oriented* and *support* activities within an ontology engineering process (cf. Fig. 1). The organizational setting of the overall process is covered by so-called ontology management activities. In the pre-development phase the *feasibility study* examines if an ontology-based application or the use of an ontology in a given context is the right way to solve the problem at hand. *Domain analysis*, *conceptualization* and *implementation* are classical ontology development activities. The *maintenance* and the *use* of the ontology are post-development activities. Ontology support activ-

---

[3] Refer for example to [7, 25] for recent overviews.

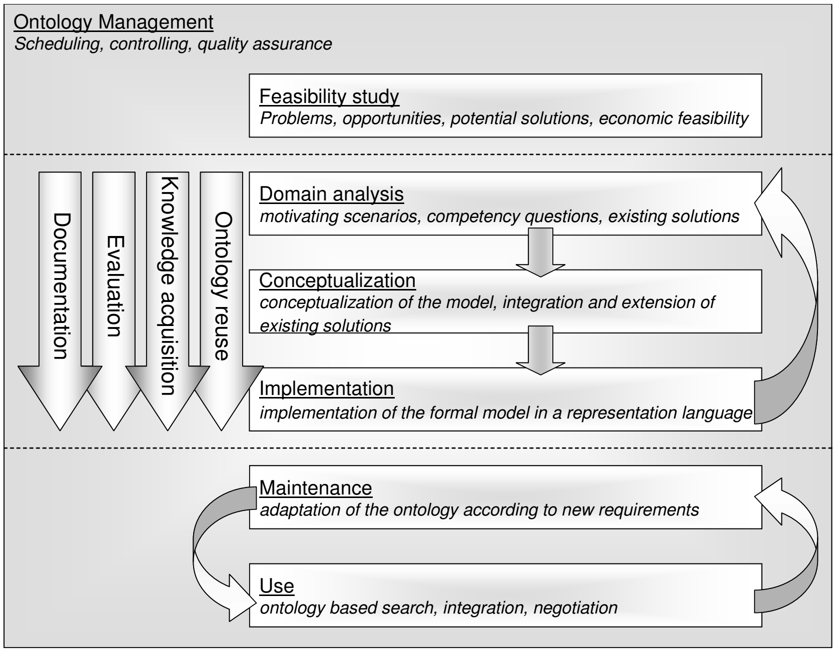


Fig. 1. Ontology Engineering Activities

ities *e.g.*, *knowledge acquisition* (KA), *evaluation*, *reuse*, and *documentation* are performed in parallel to the core development activities.

Methodologies additionally define the roles of the individuals involved in the ontology building process. They primarily differentiate between *domain experts* providing knowledge w.r.t. the domain to be modeled, *ontology engineers* with expertise in fields such as knowledge representation or ontology tools, and *users* applying the ontology for a particular purpose.

4 Our Survey

4.1 Survey Overview

The survey had the objective to capture the basic understanding of semantic technology applicants w.r.t. ontology development, to give an account of current ontology engineering practice, and to identify common problems with available ontology engineering methodologies, methods and tools.

The findings reported in this paper are based on 34 structured interviews conducted within a three months period.[4] After a short tutorial on the utilized ontology engineering terminology, the participants were requested to answer 28 questions related to particular aspects of ontology development. Complementarily to detailed answers to these questions the interviewers collected general comments.

---

[4] The detailed survey results may be obtained from the authors on request.

The survey gives a comprehensive assessment of the current state of the art in ontology engineering. Prior to the data collection procedure, the contents, organization and presentation of the survey were evaluated and revised by a group of three academia and industry experts in the area of ontology engineering. Moreover, the respondents are representative for the community of users and developers of semantic technologies. They were IT practitioners, researchers and experts from various disciplines, affiliated to industry or academia, who were involved in the last 3 to 4 years in ontology building projects in areas such as skill management, human resources, medical information systems, legal information systems, multimedia, Web services, and digital libraries. The survey was targeted *exclusively* at technology applicants (as opposed to methodology or tool developers in the Semantic Web area) in order to give a real account on the impact of the results achieved so far beyond their originating context. At the time of the interviews the interviewees possessed an average ontology engineering experience of 1 to 1.5 years. Around 50% were affiliated to industry. Only a small fraction, mostly domain experts, had received ontology engineering training in advance.

The target application the ontologies were built for ranged from proof-of-concept implementations to commercial solutions. Consequently, most of the surveyed ontologies were domain ontologies—either application-dependent or -independent. A single ontology had upper-level character, while 6 were core ontologies. The ontologies had an average size of 1000 ontology entities (concepts, properties, axioms and fixed instances). The development efforts were approximated to 4 person months on the average. 50% of the ontologies were built from scratch. If other ontologies were reused, they made up to 50% of the final ontology.

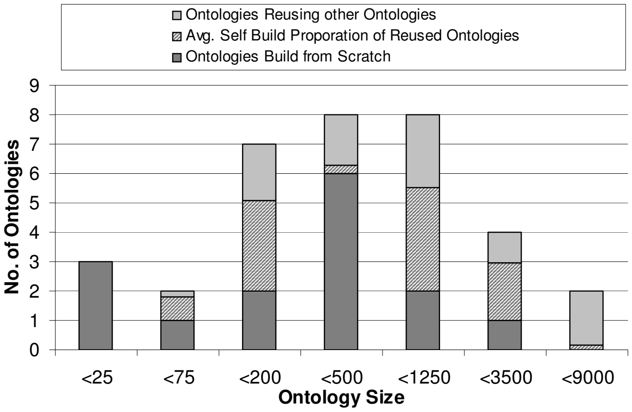


Fig. 2. Size Distribution of Surveyed Ontologies

4.2 Survey Design

The study covered both **open-ended** and **close-ended** questions (cf. Table 1).[5] The former do not impose any constraints on the form or the content of the responses, and are intended to capture general facts about the surveyed ontology engineering projects. By contrast, in the second category the answers of the respondents are limited to a fixed set of responses. Typical examples of close-ended questions are dichotomous (yes/no) questions, multiple choices, as well as scaled (also called ranking) questions using various scale models. In our case we used five point ranking scales to assess the complexity of ontology engineering activities, the quality of the methodological and tool support and the level of experience of the engineering team.

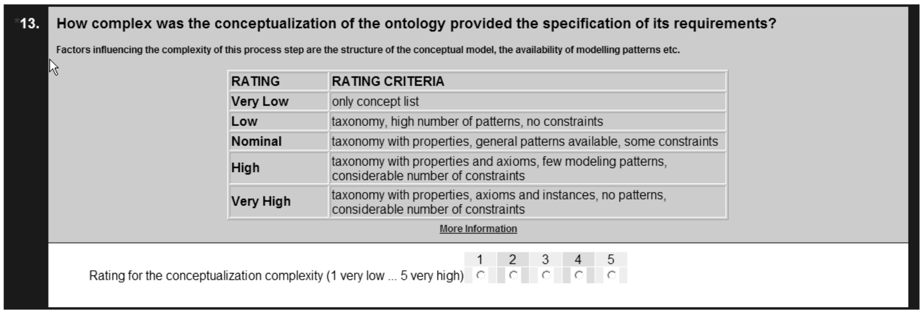
No.	**Acronym**[6]	**Topic**	**Response**
		Introductory questions	
1	ONTNAME	The name of the ontology	open-ended
2	ONTNS	The namespace of the ontology	open-ended
3	SCOPE	Purpose and scope of the ontology	open-ended
4	SIZEO	Total size of the ontology	open-ended
5	TYPE[7]	Ontology type	scaled
6	COSTS	Ontology development effort in person months	open-ended
		Process questions	
7	SIZEB	Percentage of final ontology built from scratch	open-ended
8	DCPLX	Complexity of the domain analysis	scaled
9	CCPLX	Complexity of the ontology conceptualization	scaled
10	ICPLX	Complexity of the ontology implementation	scaled
11	DATA	Complexity of the ontology instantiation	scaled
12	SIZER	Percentage of the final ontology built by reuse	open-ended
13	COMPRH	Complexity of the ontology understandability task	scaled
14	USAB	Complexity of the usability assessment	scaled
15	TRANS	Complexity of translation operations	scaled
16	MOD	Complexity of modification operations	scaled
17	INT	Complexity of merging and integration tasks	scaled
18	DOCU	Complexity of the documentation task	scaled
19	OEVAL	Complexity of the evaluation of the final ontology	scaled
		Organizational questions	
20/21	OCAP/DCAP	Capability of the ontologists/domain experts	scaled
22/23	OEXP/DEXP	Level of experience of the ontologists/domain experts	scaled
24/25	LEXP/TEXP	Level of experience w.r.t. languages and tools	scaled
26	SIZET	Team size	open-ended
27	TURN	Personnel turnover	scaled
		Technological questions	
28	TOOL	Level of technological support for particular activities	scaled

Table 1. Survey Organization

---

[5] Refer to [23] for a detailed account of questionnaire design principles.

The questionnaire can be divided into 4 categories. 6 **introductory questions** are intended to describe the most important facts about an ontology engineering project: the ontology which was built, its name, size, scope, purpose, as well as the overall development efforts. The second group of 13 questions focused explicitly on particular aspects of the ontology engineering process (**process questions**, cf. Table 1): the domain analysis, the conceptualization, the implementation, the ontology population as well as reuse and knowledge acquisition. For each ranking question the ontology builders rated on a scala from 1 to 5 whether the respective activity was very easy (1) or very difficult (5) to perform in their case. For each rating level and each question, the survey included detailed examples in order to facilitate the ranking and make the results comparable. For instance, in the case of the *conceptualization* the interviewees estimated the score in relation to the complexity of the conceptual model (cf. Figure 3). If reuse was relevant, they were asked to estimate the contingent of the final ontology, which was built in this way, and to rate the understandability and clarity of the source ontologies, the complexity of the translation between differen representation languages, the difficulties in adapting the reused ontologies to the own needs, or the integration of multiple reused ontologies. The complexity of the evaluation and documentation were further aspects of the questionnaire.



*13. **How complex was the conceptualization of the ontology provided the specification of its requirements?**

Factors influencing the complexity of this process step are the structure of the conceptual model, the availability of modelling patterns etc.

RATING	RATING CRITERIA
Very Low	only concept list
Low	taxonomy, high number of patterns, no constraints
Nominal	taxonomy with properties, general patterns available, some constraints
High	taxonomy with properties and axioms, few modeling patterns, considerable number of constraints
Very High	taxonomy with properties, axioms and instances, no patterns, considerable number of constraints

More Information

	1	2	3	4	5
Rating for the conceptualization complexity (1 very low ... 5 very high)	○	○	○	○	○

Fig. 3. The Conceptualization Complexity Question

The third part of the survey contained 8 **organizational questions**. The interviewees assessed ranking for the average capability of the ontology building team as domain experts and ontology engineers, respectively. Further questions covered, for instance, the average experience in building ontologies and the available know-how as regarding ontology representation languages and tools. A last part of the survey was concerned with **technological** issues, primarily related to the available tool and methodological support and its quality. The interviewed persons also commented on the activities which were most difficult to carry out, and on the methods and tools utilized in particular projects.

---

[5] We make use of these notations within the discussion of the survey results.

[7] The scale model for the description of ontology types included 5 elements: upper-level, domain, application, task and core ontologies cf. [8].

The survey contributed to a better understanding of the ontology building process in general. Furthermore, the results of the survey provide more detailed information about the effort related to ontology building in the future.

4.3 Survey Results

The presentation of the results of the conducted interviews is divided into four categories: *general issues*, *process issues*, *organizational issues* and *technological issues.*

General Issues This category of results is not directly related to specific survey questions, but are distilled from the plethora of general concerns and comments expressed by the interviewees during the operation of the study.

The survey clearly pointed out that the popularity of ontologies—as a means to solve many non-trivial IT problems—is not equally shared by the discipline of ontology engineering. While the reasons for this situation are traceable in the case of the interviewed domain experts, many of the IT professionals or researchers seemed not to perceive ontology building as a systematic process which should be performed according to a pre-defined methodology. Nevertheless, the way ontology development has been carried out was on the whole compatible with the recommendations made by methodologies available in the field (see below). A more serious issue is related to the lack of terminological knowledge and to the controversial understanding of the participants w.r.t. core concepts of ontology engineering. Starting from the oft-enunciated fuzzy definition of ontologies, this confusion is further propagated to activities involving the usage of external knowledge sources to aid the manual ontology development. The survey showed that the majority of the participants associate ontologies with almost every type of lightweight conceptual structure, and that they have difficulties in distinguishing between tasks such as ontology reuse, ontology learning and, more generally, knowledge acquisition. Further on, the notion of reuse was often associated with informative materials consulted by domain experts while constructing the ontology. A last terminological weak-point is constituted by the basic understanding of IT experts, who did not distinguish between a conceptual and an implementation level of an ontology. Terms such as "conceptualization" and "formalization" seemed not to be prevalent in current ontology engineering practice.

Result
80% of the projects did not follow a particular ontology engineering methodology
90% of the projects followed implicitly established ontology engineering activities
80% of the participants expressed the need for a terminological clarification
very broad understanding of the term *ontology* in 100% of the cases
ontology reuse interpreted as usage of arbitrary information sources in 90% of the relevant cases

Table 2. Summary of General Issues

Process Issues As aforementioned the structure of the survey assumed a "classical" breakdown of ontology engineering processes at the level of activities introduced in Section 3. This set-up proved to match to a satisfactory extent to the way the survey participants carried out the process. The interviews emphasized however some discrepancies between i) the complexity and significance of particular process stages as perceived by ontology engineering practitioners, and ii) the attention these process stages received in the research community so far. This applies in particular to the following issues:

Domain analysis: All participants emphasized the resource-intensive nature of this process step and the lack of low entry barrier methods and tools to support the knowledge elicitation task. Moreover, in projects building highly specialized ontologies such as for the legal or the medical domain, ontology engineers—who were responsible for guiding the rest of the team during this phase—manifested their concern w.r.t. their ability to accomplish this task appropriately. Consequently the interviewed engineers commonly agreed on the (partial) arbitrariness of the knowledge elicitation procedure in the questioned projects. Methodologies tend to handle these particular issues at a very generic level and were thus not utilized for the domain analysis, excerpt in form of competency questions in individual cases. Further on, the interviews revealed the difficulties encountered by ontology developers in combining multiple strategies for building fragments of the same ontology. Some of them were puzzled about the means to choose among alternative strategies and the way and the time point partial engineering results emerging from these complementary activities should be integrated.

Conceptualization and implementation: The majority of the interviewees did not perceive a clear cut between the conceptualization and the implementation steps as necessary. After a lightweight description and classification of the expected outcomes the engineering team implemented the ontology with the help of a common ontology editor. This task was primarily performed by domain experts, who did not report any particular difficulties in getting familiar with or utilizing simple ontology editors.[8]

Ontology reuse: In over 50% of the cases the final ontology was built with the help of other ontological sources. In 11 of the reuse-relevant projects the engineering team pointed out the considerable resources invested in understanding the reused sources and assessed high complexity ratings to the task of translation. By contrast, issues like merging and integration were not relevant to the investigated scenarios or were assigned lower complexity scores. This situation might, however, be caused by the fact that the average number of reused ontologies per project was very low (maximal 2). The modification of the ontologies to be reused was performed in two stages: first the relevant fragments were extracted from the sources, then they were customized and integrated to the target ontology. The effort required to perform these activities was often estimated to a nominal value.

---

[8] This situation might correlate with the fact that of the ontologies had a relatively simple structure, i.e. taxonomies augmented by properties between concepts.

Ontology evaluation: All domain experts manifested concerns w.r.t. the quality of the implemented ontology. They claimed for usable methods guiding them within this task. This holds true for both ontology engineering experts, who were not able to appropriately assist this endeavor, but also for the non-IT survey participants, and caused acceptance problems. In over 40% of the cases the results of the process were not nominally evaluated. In the remaining ones, 95% have been manually evaluated by their authors without the help of a methodical approach. 3 ontologies have been evaluated through external expert judgement. A further result of the study was the necessity of an incremental approach to ontology evaluation. Ontology developers were not sure about the most adequate point in time for performing first evaluation tests, and pointed out the implications of this issue for evolving ontologies or for long-term projects, in which the application using the created ontology is not timely available for preliminary evaluations. In this context they spoken out the question of how to feasibly determine the real start and the end points of an ontology engineering project, this being a prerequisite for any controlling and planing activities.

Ontology population: The question related to the complexity of the ontology population task addressed two aspects: the data sources used as input to create the ontology instances and the required mapping between the input scheme and the target ontological model. Both were estimated with the highest scores among all questions on ontology development (between 3 and 4), comparable to the ones associated to the domain analysis. In most of the cases the ontology population task was exercised on semi-structured data expressed in natural language and required complex mappings covering concepts and properties.

Ontology maintenance: This task was not properly represented in any of the analyzed projects. Some ontology engineers raised, however, the issue of clearly determining the transition point from ontology development to maintenance or evolution. In correlation with the fuzzy nature of the evaluation task, ontology engineers were not in the position to distinguish between the two phases.

Documentation: The majority of the projects reported above average documentation efforts, and expressed the need for an automatization of this task.

The survey pointed out that only a small percentage of ontology-related projects follow a systematic approach to ontology building, and even less commit to a specific methodology. Most of the projects are executed in an ad-hoc manner. In the early project phases ontology builders underestimate the efforts underlying this endeavor and the importance of a methodological framework. It is the early project phase, however, where methodologies could offer most gains, as our analysis illustrates. In later stages of the process this perception changes; however, the a posteriori adoption of a certain development strategy is complicated by the lack of methods to evaluate the suitability and the general quality of existing methodologies.

The activities carried out in the examined case studies are covered by at least one methodology. However, the provided guidelines are insufficient for the most challenging process stages and no methodology handles the complete range of activities registered in our survey. Furthermore, different real-world scenarios require customizable method assemblies, rather than pre-defined rigid workflows as proposed by current methodologies.

Question acronym	**Average result**[9]	**Comments**
DCPLX	3.3	lack of fine-grained guidance for the domain analysis resource-intensive activity challenging if the ontology integrates knowledge from diff. domains
CCPLX	3	widespread use of ontology editors no clear distinction between conceptualization and implementation
ICPLX	3.5	use of existing structured data sources to populate ontology extraction of instances from text
SIZER	50%	max. 2 ontologies reused for a new ontology 50% of the final ontology built by reuse
TRANS	3.5	high complexity of language translation
MOD	3.5	mainly partial reuse involving customization
OEVAL	3	need for clear guidelines for ontology evaluation very complex activity, rarely performed
DOCU	3.5	documentation is a time consuming activity dedicated tool support required

Additional comments
lack of guidelines for the use of ontology learning algorithms lack of guidelines to combine different knowledge acquisition techniques no methodology covers all relevant ontology engineering activities need for methodological support seldom perceived at project start

Table 3. Summary of Process Issues

Organizational Issues The surveyed case studies contained on the average relatively small teams (two ontology engineers and two domain experts), which regularly organized F2F meetings to help on the ontology development. The personnel turnover was very low, this being also related to the relatively short project duration. In this context we can not make any reliable statement on the effects of this factor in long-standing projects. Nevertheless, in the projects in which domain experts did not possess IT competency the participants reported communication and comprehension problems. Prior to specifying the ontology, considerable efforts were invested in agreeing on a similar understanding upon the domain of interest. This issue had consequences on the way the engineering process was performed. However, we did not record major problems w.r.t. the

---

[9] We used the following five point scale 1:very low, 2:low, 3:nominal, 4:high, 5:very high.

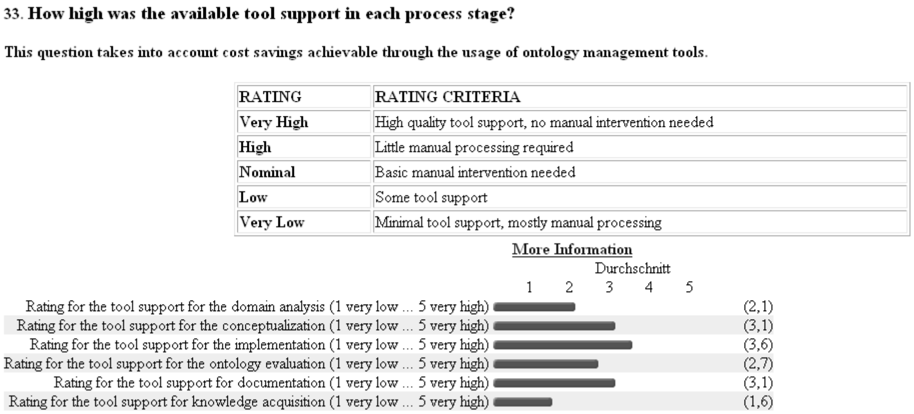
achievement of a shared ontology, though the importance of methods assisting the consensus making process was acknowledged by many interviewees.

Question acronym	**Average result**
SIZET	4
OEXP/DEXP	1.25 years
TURN	15% personal turnover

Additional comments
definition of common terminology in teams often difficult
process support for consensus making required

Table 4. Summary of Organizational Issues

Technological Issues As stated in other similar investigations in the past, tools are a crucial factor for an efficient ontology engineering practice. As illustrated in Figure 4, phases such as the conceptualization and the implementation received satisfactory ratings, while the evaluation and the domain analysis are de facto manual tasks.



33. **How high was the available tool support in each process stage?**

This question takes into account cost savings achievable through the usage of ontology management tools.

RATING	RATING CRITERIA
Very High	High quality tool support, no manual intervention needed
High	Little manual processing required
Nominal	Basic manual intervention needed
Low	Some tool support
Very Low	Minimal tool support, mostly manual processing

Rating for the tool support for the domain analysis (1 very low ... 5 very high)	(2,1)
Rating for the tool support for the conceptualization (1 very low ... 5 very high)	(3,1)
Rating for the tool support for the implementation (1 very low ... 5 very high)	(3,6)
Rating for the tool support for the ontology evaluation (1 very low ... 5 very high)	(2,7)
Rating for the tool support for documentation (1 very low ... 5 very high)	(3,1)
Rating for the tool support for knowledge acquisition (1 very low ... 5 very high)	(1,6)

Fig. 4. The Results of the Tool Support Question

Besides the well known absence of computer-aided ontology evaluation support, our survey pointed out the need for dedicated conceptualization tools. All investigated projects used ontology editors for this purpose. Their functionality was positively evaluated by many of the participants. Further on, the answers repeatedly referred to the need for tools for translating between representation languages, including methods to convert semi-structured conceptual structures to OWL and RDF(S). The notion of lightweight technological support was mentioned in relation to every ontology engineering activity: apart from the challenging research questions approached in the last fifteen years, the community of ontology engineering practitioners require simple means to extract ontological structures from existing knowledge sources, to translate concept labels to other natural languages and to ease the creation of documentation.

Question acronym	**Average result**	**Comments**
TOOL: domain analysis	2,1	technological support limited to text analysis tools
TOOL: conceptualization	3,1	satisfactory support for conceptualization conceptualization supported by text editors, mind maps and OE environments
TOOL: implementation	3,6	satisfactory results for OE environments no lightweight ontology engineering environments available no tool supports the easy translation of an ontology to different natural languages
TOOL: evaluation	2,7	ontology editors used for technical evaluation of ontologies semantic ontology evaluation not tool supported
TOOL: documentation	3,1	text editors no specialized ontology documentation tools available
TOOL: knowledge acquisition	1,6	existing KA tools not used or not helpful in most cases no tools to leverage ontologies from existing data sources

Table 5. Summary of Technological Issues

5 Discussions of the Results

The survey highlighted several weak-points of existing ontology engineering methodologies from an empirical point of view. In this section we introduce research directions which explicitly address the problems the community of ontology engineering practitioners is facing, and sketch a possible solution for the alleviation of the present state of the art.

General Issues Despite the long history and the multitude of research initiatives in the area of ontology engineering methodologies, the visibility of the achieved results remains restricted to a small community of experts affiliated to academia. Methodologists should thus invest more efforts in the dissemination of their results to a wider audience and in promoting the achievable returns of a methodologically supported ontology building approach. Case study reports comparing ontology building efforts with and without methodological support may be one way to demonstrate the efficiency gains. Further on there is a need for support in selecting methodologies which suit certain application settings. The advantages and disadvantages of manual, semi-automatic or reuse-oriented ontology creation for a specific use case are not obvious for potential ontology builders, even when they have a strong IT background. A methodology selection framework should include application-oriented decision criteria to support this activity. Moreover, the methodology itself could include a step in which the engineers pick from a list of available methods the ones suitable for a particular task and build up their own process model. Template process models covering standard requirements could be made available. This is an alternative to the current trend to create new methodologies for emerging application scenarios. Once the workflow has been specified, the developers need requirements engineering support techniques. This includes, for instance, means to determine the optimal formality level for the prospected ontology or to analyze the trade-off between development effort and size/complexity of the conceptual model.

Process Issues Knowledge elicitation is one of the most time consuming tasks in the process of building an ontology. Current methodologies lack a comprehensive guidance for this task and instruments to choose among knowledge elicitation techniques. With the introduction of selection metrics this problem may be alleviate. A comprehensive methodology should allow for modular and customizable process models in which ontology builders combine different methods for the stages of the ontology development process they require. Ontology builders may leave out process descriptions for activities which are automated by tools. Moreover, methodologies should provide more support on the evaluation of ontologies beyond completeness and soundness. In commercial applications trade-offs between additional modeling and costs are balanced. This requires an evaluation w.r.t. the achievable gains of modeling additional ontology entities.

No.	**Recommendation**
	General recommendations
1	enforce dissemination, *e.g.* publish more best practices
2	define selection criteria for methodologies
3	define a unified methodology following a method engineering approach
4	support decision for the appropriate formality level given a specific use case
	Process recommendations
5	define selection criteria for different KA techniques
6	introduce process description for the application of different KA techniques
7	improve documentation of existing ontologies
8	improve ontology location facilities
9	build robust translators between formalisms
10	build modular ontologies
11	define metrics for ontology evaluation
12	offer user oriented process descriptions for ontology evaluation
	Organizational recommendations
13	provide ontology engineering activity descriptions using domain-specific terminology
14	improve consensus making process support
	Technological recommendations
15	provide tools to extract ontologies from structured data sources
16	build light-weight ontology engineering environments
17	improve the quality of tools for domain analysis, ontology evaluation, documentation
18	include methodological support in ontology editors
19	build tools supporting collaborative ontology engineering

Table 6. Summary of Recommendations

Organizational Issues Ontology builders had no major organizational problems. For collaborative ontology engineering in distributed environments with participants originating from different domains, methodologies should offer support on the consensus finding process. The reformulation of existing method descriptions in a less computer science-oriented way, using domain-specific terminology, may facilitate the comprehensibility of the engineering process. Positive experiences in the biology domain attest the value of such domain-close descriptions.

Technological Issues From a technological point of view several ontology management tools (e.g. editors or reasoners) have reached a feasible maturity level. Nevertheless many ontology engineering activities are not supported adequately at technical level. Available tools in this context originate from academic research projects, and focus on solving non-trivial generic research questions instead of operationalizing simple tasks whose automatization is clearly not problematic. Further on ontology builders require easy-to-use ontology engineering environments which are extensible to support different kinds of ontology engineering processes. Current editors do not support discussion-based ontology engineering, in which a number of ontology builders first argue about modeling decisions before they decide on them. The provision of ontology engineering patterns and high-quality ontologies for commonly used domains may also ease the effort to build ontologies.

In summary, this survey demonstrates the limited impact of methodologies in real-world ontology-related projects. However, the results also evidence that ontology engineering research has reached a level of development with a basic inventory of methods and tools which, if properly utilized, considerably ease the work of ontology practitioners. Therefore, a first conclusion of the survey is the need for initiatives aiming at promoting these results to a wider audience and consolidating them for an increased usability. European projects such as Sekt-Semantically-Enabled Knowledge Technologies are taking first steps in this direction. Complementarily, the experiences gained during this survey let us assume that a *method engineering* approach to the area of ontologies could be a viable alternative to the creation of new methodologies and methods. Software engineering is already moving in this direction, *cf.* [9]. They describe software engineering activities according to predefined templates. Templates include input and output factors, available methods to support the activity, required preceding activities and possible succeeding activities among other things. Additional to this new methodological approach practitioners require evaluation methods to compare the trade-offs between investing in additional modeling and gained functionality.

6 Summary and Outlook

Despite the growing popularity of ontologies as a knowledge representation formalism used on the Web, very little is known about the engineering process underlying their construction in practice. The literature reports predominately on case studies which involved methodologists, while ontologies are envisioned to be built by domain experts possessing limited to no professional skills in ontology engineering. We alleviate this information gap with a survey interviewing over 30 practitioners who developed ontologies for commercial, as well as academic applications for a wide range of domains. Although this number is not large w.r.t. the number of ontologies available on the Web it is the largest study conducted so far; and we continue to collect ontology engineering experiences.

The survey investigated the systematics, the invested effort and the general problems encountered in building these ontologies. The main findings are i) practitioners do not follow any particular ontology engineering methodology, though there is some overlapping ii) they require selection support to choose from manual, semi-automatic or reuse oriented engineering approaches, iii) they need cost benefit analysis methods to determine the transition point between ontology engineering activities and iv) existing ontology management tools have reached a feasible level of functionality to be useful.

In order to overcome some of the problems revealed in this survey we suggest to increase the efforts invested in promoting the advantages of methodological ontology engineering to a wider audience. Furthermore, we propose to establish a unified methodology which supports domain experts to customize ontology engineering process models according to their application scenario. This requires the creation of compliant method components or the adjustment of existing ones to enable their joint utilization. Evaluation metrics for ontologies need to be defined in order to assess the usability of similarly scoped methods in particular circumstances. We will continue our research in these directions.

Acknowledgements This work has been partially supported by the European Network of Excellence “KnowledgeWeb-Realizing the Semantic Web” (FP6-507482), as part of the KnowledgeWeb researcher exchange program **T-REX**, and by the European project “Sekt-Semantically-Enabled Knowledge Technologies”(EU IST IP 2003-506826). We thank all interviewees for the valuable input without which this paper could not have been produced. We also want to thank York Sure (University of Karlsruhe) and Richard Benjamins (iSOCO) for contributing to the realization of the final version of the survey.

References

1. V. R. Benjamins, D. Fensel, S. Decker, and A. Gómez-Pérez. $(KA)^2$: Building Ontologies for the Internet. *International Journal of Human-Computer Studies*, 51(1):687–712, 1999.
2. A. Bernaras, I. Laresgoiti, and J. Corera. Building and Reusing Ontologies for Electrical Network Applications. In *European Confernce on Artificial Intelligence (ECAI'96)*, 1996.
3. O. Corcho, M. Fernández-López, and A. Gómez-Pérez. Methodologies, tools and languages for building ontologies: where is their meeting point? *Data & Knowledge Engineering*, 46(1):41–64, 2003.
4. J. Euzenat. Building Consensual Knowledge Bases: Context and Architecture. In *Proc. of the 2nd Int. Conference on Building and Sharing Very Large-Scale Knowledge Bases (KBKS)*, pages 143–155, Enschede the Netherlands, 1995.

5. M. Fernández-López, A. Gómez-Pérez, J. P. Sierra, and A. P. Sierra. Building a Chemical Ontology Using Methontology and the Ontology Design Environment. *Intelligent Systems*, 14(1), January/February 1999.
6. A. Gangemi, D.M. Pisanelli, and G. Steve. Ontology Integration: Experiences with Medical Terminologies. In *Formal Ontology in Information Systems*. IOS Press, 1998.
7. A. Gómez-Pérez, M. Fernández-López, and O. Corcho. *Ontological Engineering*. Springer, 2003.
8. N. Guarino. Formal Ontology and Information Systems. In *Proc. of the 1st Int. Conf. on Formal Ontologies in Information Systems (FOIS)*, 1998.
9. B. Henderson-Sellers. Method engineering for OO system development. *Communications of the ACM*, 46(10):73–78, 2003.
10. M. Hristozova and L. Sterling. Experiences with Ontology Development for Value-Added Publishing. In *Proceedings of the International Workshop on Ontologies in Agent Systems OAS03 co-located with the AAMAS*, 2003.
11. D. Jones, T. Bench-Capon, and P. Visser. Methodologies for Ontology Development. In *Proc of the IT&KNOWS Conference of the 15th IFIP World Computer Congress*, 1998.
12. N. J. J. P. Koenderink, J. L. Top, and L. J. van Vliet. Expert-Based Ontology Construction: A Case-Study in Horticulture. In *Proceedings of the International Database and Expert Systems Applications Workshops DEXA05*, 2005.
13. I. Laresgoiti, A. Anjewierden, A. Bernaras, J. Corera, Schreiber A. T., and B. J. Wielinga. Ontologies as vehicles for reuse: a mini-experiment. In *Proceedings of the 10th Banff Knowledge Acquisition for Knowledge-Based Systems Workshop K-CAP96*, pages 1–21, 1996.
14. T. Lau and Y. Sure. Introducing Ontology-based Skills Management at a large Insurance Company. In *Modellierung 2002, Modellierung in der Praxis - Modellierung für die Praxis, Tutzing, Deutschland, 25.-27. März 2002*, pages 123–134, 2002.
15. D. B. Lenat and R. V. Guha. *Building large knowledge-based systems. Representation and inference in the CYC project*. Addison-Wesley, Reading, Massachusetts, 1990.
16. A. Maedche. *Ontology Learning for the Semantic Web*. Kluwer Academics, February 2002.
17. E. Paslaru Bontas. Practical Experiences in Building Ontology-Based Retrieval Systems. In *Proceedings of the 1st International ISWC Workshop on Semantic Web Case Studies and Best Practices for eBusiness SWCASE05*, 2005.
18. E. Paslaru Bontas, M. Mochol, and R. Tolksdorf. Case Studies in Ontology Reuse. In *Proc. of the 5th International Conference on Knowledge Management IKNOW05*, 2005.
19. H. S. Pinto and J. Martins. Reusing ontologies. In *AAAI 2000 Spring Symposium on Bringing Knowledge to Business Processes*, pages 77–84, 2000.
20. H. S. Pinto, C. Tempich, and S. Staab. DILIGENT: Towards a fine-grained methodology for DIstributed, Loosely-controlled and evolvInG Engingeering of oNTologies. In *Proc. of the 16th European Conference on Artificial Intelligence (ECAI 2004)*, pages 393–397, 2004.
21. Dip European Project. Financial Ontology (Deliverable 10.3 DIP FP6 - 507483). http://dip.semanticweb.org/documents/D10.3.pdf, 2005.
22. T. Russ, A. Valente, R. MacGregor, and W. Swartout. Practical Experiences in Trading Off Ontology Usability and Reusability. In *Proc. of the Knowledge Acquisition Workshop KAW99*, 1999.

23. Seymour Sudman and Norman M. Bradburn. *Asking Questions : A Practical Guide to Questionnaire Design*. Jossey-Bass, 1st edition, 1982.
24. Y. Sure, S. Bloehdorn, P. Haase, J. Hartmann, and D. Oberle. The SWRC Ontology - Semantic Web for Research Communities. In *Proc. of the 12th Portuguese Conference on Artificial Intelligence - Progress in Artificial Intelligence (EPIA 2005)*, pages 218 – 231, 2005.
25. Y. Sure, C. Tempich, and D. Vrandecic. Ontology Engineering Methodologies. In *Semantic Web Technologies: Trends and Research in Ontology-based Systems*. Wiley, UK, 2006.
26. C. Tautz and K. D. Althoff. A Case Study on Engineering Ontologies and Related Processes for Sharing Software Engineering Experience. In *Proc. of the International Conference on Software Engineering and Knowledge Engineering SEKE00*, 2000.
27. C. Tempich, H. S. Pinto, and S. Staab. Ontology Engineering Revisited: an Iterative Case Study with DILIGENT. In *Proc. of the 3rd European Semantic Web Conference (ESWC 2006)*, pages 110–124, 2006.
28. M. Uschold, M. Healy, K. Williamson, P. Clark, and S. Woods. Ontology Reuse and Application. In *Int. Conference on Formal Ontology and Information Systems FOIS98*, 1998.
29. M. Uschold, M. King, S. Moralee, and Y. Zorgios. The Enterprise Ontology. *Knowledge Engineering Review*, 13(1):31–89, 1998.
30. G. Zhao, J. Kingston, K. Kerremans, F. Coppens, R. Verlinden, R. Temmerman, and R. Meersman. Towards an Ontology of Forensics covering Financial Securities Fraud. In *Proceedings of the International Workshop on Regulatory Ontologies WORM04 co-located with the OTM Conferences*, 2004.