Towards a New Synthesis of Ontology Technology and Knowledge Management

Technical Report IR-BI-001 [1]

Peter Mika and Hans Akkermans
Free University Amsterdam VUA
Faculty of Sciences, Section Business Informatics
De Boelelaan 1081a, 1081HV Amsterdam, The Netherlands

March, 2004

---

[1] Submitted to the Knowledge Engineering Review in April, 2004.

Abstract

This paper provides a synthesis of ontology research and one of its most significant application areas, the business discipline of Knowledge Management (KM for short). Despite that such a synergy is in great demand between consumers and suppliers of ontology technology, there is a lack of understanding on one side and a lack of realism on the other. The lack of communication and understanding between ontology research and applied Knowledge Management holds the danger that mainstream KM will remain neutral to ontologies, while research develops as a largely unguided exploration.

We provide an analysis of the state-of-the-art in technologies for Knowledge Management in order to explain the role and potential contribution of ontologies to the development of Knowledge Management. Our main contribution towards a synthesis of Ontology Technology and Knowledge Management is a visual classification framework called the Semantic Web Matrix, which gives an easily understood, business-oriented account of application scenarios for ontology-based Knowledge Management, both centralized and distributed. We conclude our survey by discussing the most significant bottleneck in applying ontologies, namely a lack of understanding regarding the social nature of knowledge and the consequences this has for efforts targeted at the management of distributed knowledge.

1 Introduction

In the last decade knowledge has been recognized as a key success factor for both businesses and societies at large [Drucker, 1993; Nonaka and Takeuchi, 1995; Davenport and Prusak, 1998].

While companies have been managing their human and intellectual assets long before, the business discipline of Knowledge Management (KM for short) was eventually formed to meet these specific challenges. Unfortunately, there is hardly any consensus on the definition of Knowledge Management, mostly due to the varied views of the practitioners on what knowledge is on the first place. As a result, sometimes even the concept of Knowledge Management and its validity as a discipline comes under attack, see e.g. [Wilson, 2002]. Although there is no dispute concerning the importance of corporate knowledge as such, it is sometimes questioned to what extent management of (often tacit) knowledge is possible at all, and what the added value of specific KM ideas is.

For our purposes, we will define Knowledge Management as the collection of measures aimed at maintaining a continuous flow of knowledge in such a way that the appropriate knowledge reaches the right members of the organization at the right time and in an actionable format. This allows organizations to learn more effectively as a community and to maximize the returns on their individual and collective information sources and problem-solving capabilities. KM achieves this goal by providing the necessary structures, processes and technology for collecting, retaining[1], analyzing and interpreting, sharing, using and maintaining business critical knowledge. In its methods, Knowledge Management builds on the rapid convergence of ICT, business process and workflow management, Project Management, Human Resource Management (HRM), organizational behaviour, communication and library sciences.

At the same time that Knowledge Management emerged and took hold in the business field, major research efforts have been focused in the Artificial Intelligence community on the issues of Knowledge Representation (KR) and Reasoning, combining earlier experiences in conceptual modelling with formal logics. This work centers around the notion of ontologies, which represent a shared, formal understanding of a domain theory [Gruber, 1993; Borst *et al.*, 1997]. The term *shared* refers to an agreement within a community of experts over the description of their domain, while *formal* indicates the representation of this agreement in some sort of computer-understandable format. (Note also the rather vague notion of a domain theory in the definition: ontology research makes no claim about the nature of the knowledge to be modelled.) Much interest in this work has been inspired by the idea of a Semantic Web that would apply knowledge representation techniques in a distributed, web-based scenario (see [Berners-Lee *et al.*, 2001]). We return to ontologies and the Semantic Web in Section 6.

In the ontology research community, Knowledge Management is widely acknowledged as one of the most prominent applications of ontologies [Leger, 2002].[2] Nev-

---

[1] Retaining knowledge of employees leaving an organization is one of the priorities at times when organizations engage in downsizing or simply loose employees that retire. In this case the knowledge should flow from leaving personnel to entering/remaining employees.

[2] The OntoWeb survey of business scenarios mentions e-business, information retrieval and portals as other major application areas for ontologies. However, we consider both portals and information retrieval as supporting technology and not applications on their own. This leaves KM and e-commerce as the two major

ertheless ontologies have yet to found their way into the practice of Knowledge Management in the business field, which is not to say that the need for ontologies would be lacking. In fact, as we will show in Section 3, state-of-the-art ‘intelligent’ knowledge management solutions have independently developed to a stage where there is a clear potential for more advanced techniques of knowledge representation. Even the word ‘ontology’ has yet to find its way into the KM vocabulary. Given the lack of understanding, mainstream KM and ontology research evolve in parallel without exploiting the synergies between application area and supporting technology.

We believe that a poor understanding of the role and potential of ontologies is the major culprit behind the sluggish adoption of ontology-based tools and technologies within the mainstream KM community. This problem not only concerns Semantic Web startups targeting the KM market, but is also a significant issue for research policy. Without analyzing this relationship and catering to its needs the ontology community risks that the benefits of the technology will not be understood by the members of the KM community and in turn ontology research will be less compelled to meet the genuine needs of KM.

What could be done to make the benefits of technology transparent? Based on a thorough analysis of the KM field in Sections 2.1 and 3, this paper takes up the challenge by revisiting the fundamental motivation for applying ontology-based methods to Knowledge Management. From this motivation we create the Semantic Web Matrix, a novel classification of application scenarios for ontology-based Knowledge Management. The primary axis of this framework is ontology functionality and usage scenarios. In Section 6 this classification is extended with a web dimension to represent the various stages of decentralization in ontology-based applications.

The Semantic Web Matrix comprises scenarios close to the efforts of applied KM and is designed to be understood in a non-technical, business environment as well as to cover all existing prototypes of ontology-based KM systems. The framework is chiefly intended to help the KM community understand ontology functionality, but should also support the ontology and Semantic Web research community to explain the existing prototype applications and discuss the evolutionary development of the technology.

Lastly, in Chapter 7 we will use our new understanding to explain the limitations of ontology-based technology in the KM area and establish future work for the research community.

2 Ontology-based methods in Knowledge Management

The business of management, as per Peter Drucker, lies in creating systems that are underpinned by strategic and tactical answers to the two questions - what is our business and what should it be?

In the era of a networked, global, knowledge-driven economy, managers and business consultants have vested significant interest in the emerging field of Knowledge Management. Even though the term KM brought hardly any new meaning and is at best a container concept for a wide array of methods, executives’ interest was underpinned by the strategic vision of an (increasingly distributed) organization that enhances its performance by continuous learning, interactions and an efficient reuse of its knowl-

---

consumers of ontology technology.

edge.

While winning the intellectual argument for KM is easy, arguing for any kind of technology poses certain challenges. Managers wary of technology undoubtedly ask: “So tell me what does all this mumbo jumbo do to help our company manage the knowledge that it needs in order to be more efficient?” At times when business pressure is growing on KM groups to make costs and benefits transparent, it is of utmost importance to come up with a cogent elevator ride answer to such a bottomline question.

Note that it is unrealistic to expect an exact measure on the return of investment (ROI) for KM in general. It is close to impossible, for example, to quantify the value of time, money, and reputations being lost over decisions that rely on outdated, incorrect or incomplete information. Similarly, one cannot provide a numerical value for the benefits of having an organization where employees truly understand the organization’s strategy, vision and goals. Most executives, however, understand this dilemma (because it is not unique to KM but has to be dealt with in any area where intangibles play a significant role, including service, marketing, branding, and public relations), and might even be wary of fictitious calculations.

Measuring the impact of technology is equally challenging, especially because technology is only a part of KM solutions (see section below.) For example, a 3% increase in the precision of document retrieval might make all the difference in the efficiency of a particular process in a certain organization, while it might fail to have any tangible effect in another setting - for example, if the knowledge bottleneck turns out to be elsewhere.

We are not looking here for a single *killer application* either. Social studies of science show that only in a few historical cases of technological innovation — including disruptive forms of innovation — this notion is meaningful at all to explain progress in adoption and diffusion (already the invention of the steam engine and of the computer are counterexamples). Such an approach also would contradict best practice experience of knowledge engineers manifested in methodologies such as CommonKADS [Schreiber *et al.*, 1999]. In CommonKADS the process of engineering starts with a carefully selected list of problems or opportunities encountered by the enterprise. It is only after that this list is filtered and ranked by feasibility (both from technical, project management and organizational points of view) that actual proposals are made for the construction of a knowledge-based system using a particular technology. Following this knowledge engineering wisdom, we will talk about *scenarios* where ontology-based methods play an enabling role or might significantly increase the profitability of applications.

In this section we first take a general look at the role of technology in KM and the present state-of-the-art. We also survey the major streams of KM; understanding the views and methods of the various schools is of utmost importance for those who market any kind of technology in this area.

Secondly, we turn to the specific contribution of ontologies and provide a classification of ontology-based applications along two dimensions: the function ontologies play in applications and the structure of the applications. We aim to clarify here the added value of ontologies compared to more conventional, readily available technology, which are the baseline in a business context. We do so, as we often see non-technical onlookers first facing a semantic portal and saying“Hey, this is just an expensive website!” These kinds of statements reflect the collision between the intractable nature of the technology and the sometimes exaggerated expectations of the general audience. Therefore, it is critical to provide a contrast to well-understood, traditional (and certainly less expensive) technologies and use this as a basis for arguments about next-generation KM solutions that truly leverage ontologies.

2.1 Streams of Knowledge Management

Today there is a general agreement that KM is a complex mix of issues regarding people, organizations and technologies. Understanding this complexity is important for developing technical solutions that are at least aware of other aspects of the discipline.

Beyond the consensus over the need for a balance, the Knowledge Management community is deeply divided over the kind of technology to be used and the exact role it should play in its methods. This division reflects the variety of views over the nature of knowledge itself and the ensuing discussions about the agenda of Knowledge Management.

In the following, we will describe some of the major streams or schools of thought in KM to illustrate the issues under discussion. We have to note that many variations exist and in fact most KM practitioners do not take these opinions to the extreme.

2.1.1 Knowledge as a resource vs. knowledge as process

One of the major divisions we find is among the ‘stock’ and ‘flow’ approaches to KM [van Engers, 2001; Abecker and van Elst, 2003]. Starting point of the ‘stock’ approach is that knowledge is an asset or product that can exist separated from the human mind and can be represented and managed as any other object or resource. Extremists of this view, called ‘instrumentalists’, believe that all or most knowledge productivity will increase provided the right information systems are implemented. According to the ‘flow’ view of knowledge, knowledge is not an objectively transferable asset, since the actors involved in the process add subjective value. This insight is based on the experience that knowledge is always interpreted in some kind of context, both personal and social. In their view, knowledge is not transferred but *recreated* when information reaches the human mind and subsequently adapts to our mental models, which comprise our culture, expectations, previous experience etc. In other words, knowledge transfer is embedded in the process of communication and understanding. Taken to the extreme, some advocates claim knowledge is nothing other than a dynamic reflection on information within a certain context. As [Lueg, 2002] explains:

> “One of the lessons learned in the context of expert systems is that it’s hardly possible to ‘capture’ knowledge as ‘knowledge can only be created dynamically in time’ [Newell, 1982]. The related work suggests that knowledge is best conceptualized as an observer-relative attribution: an agent attributes knowledge to an agent observed in order to explain the observed agent’s behavior. It’s hardly possible to find out whether the observed agent actually has knowledge as knowledge is dynamically created.”

Somewhere between the stock and flow views lies the work of Knowledge Engineers (often with an AI background) who concede that some knowledge is always tacit, in other words hidden and internalized [Nonaka and Takeuchi, 1995]. Deep knowledge (the understanding of complex rules) and ‘wisdom’ (the actionable understanding of meta-level principles), which characterize experts often take a tacit form. Knowledge Engineers contend, however, that all other forms of implicit knowledge can be elicited (acquired) by using the right techniques. Their experience shows that in many cases this is just enough to be effectively used in knowledge-based systems to automate complex human tasks. Note also that deep knowledge in an explicit form usually appears to be shallower than it actually is, compared to its mysterious, unexplained tacit form.

2.1.2 Individual vs. organizational/social level

Within KM further distinctions can be made based on the focus of practitioners, namely an emphasis on the individual vs. the organizational/social level. The ‘managerial school’ in KM believes that knowledge can be controlled by managing individuals, the so-called ‘human resources’. They put their trust in hiring inquisitive, helpful, energetic people, encouraging learning and development, and creating varied and flexible environments where people can align personal and corporate needs. Managerials manage knowledge by allocating people so that they maximize the overall productivity of their knowledge. Work and organizational psychologists, cognitive scientists provide the background work by studying phenomena in workplace behavior and learning, respectively. Outcomes of their work help to understand some very human aspects of knowledge, such as why people sometimes attribute much power and self-confidence to knowledge and become reluctant to share.

Structuralists and re-engineers, on the other hand, are examples of practitioners with an organizational focus. Members of these clusters believe that in order to improve knowledge productivity we need to analyze and transform organizations themselves. This redesign and knowledge-aware management of organizations may effect the organizational structure, business processes, supporting systems, human policies, benchmarks etc.

On the other end of the spectrum, there is a ‘liberal’ school charging that management, regardless of the intent, can lead to the death of creativity. According to them, people should be given freedom and support for developing and organizing themselves as they see fit, instead of constraining them through strong directives. On the social level, this laissez-faire approach to KM lies behind cultivating Communities of Practice/Interest.[3] CoPs are the kind of informal communities that naturally form at workplaces and are mostly recognizable by observing how members support each other in problem-solving. CoPs are a particularly valuable source of knowledge, even if they hardly ever follow organizational lines; supporting knowledge sharing and learning by providing space to them is now considered a highly effective way to improve knowledge transfer, both implicit and explicit.

In sum, there are very different schools of thought in KM considered as a business discipline and practice. They stem from different views on the nature of knowledge and its manageability. They are not directly related to issues of technology, but do tend

---

[3] According to Manville and Foote [Manville and Foote, 1996], “ … a group of professionals informally bound to one another through exposure to a common class of problems, common pursuit of solutions, and thereby themselves embodying a store of knowledge”.

to lead to different intuitions as to what the importance and role of technology is or might be. We purposely use the weak term 'intuition' here, because in our opinion the relation between KM and technology is only superficially developed in the business-oriented side of KM. There is a clear role here for technology and ontology experts to develop more solid ideas and policies.

3 Technology in Knowledge Management

Knowledge Management has been the target of much technology in the past. In the first years of the discipline, technology was touted as a panacea for all knowledge needs as vendors hastily tried to relabel information management products (project databases, best practices databases, Lotus Notes installations etc.) as Knowledge Management solutions. The second wave of IT penetration into this market saw the advancement of massive database integration and business intelligence projects. The third, current phase focuses on putting information at the fingertips of users, employees and customers alike, via intranets, portals etc. Software to provide technology underpinning the process and people aspects of KM, such as workflow applications, collaborative work spaces, project management tools etc. also appeared.

Today, we find hundreds of products and services designed to support one or the other aspect of Knowledge Management[4]. Despite advances toward ever more sophisticated methods in supporting the management of knowledge assets and processes, we also have to record that there is a sense of disillusionment that resulted from the misuse of technology in the past. KM projects failed when technology was blindly applied without any regard to the greater business context (the people, culture, process and structural issues) and technology became the culprit. Through these failures, however, experts slowly learned that KM is inherently more than deploying software. (As summed up by one KM expert, "If technology solves your problem yours was not a knowledge-management problem.") This wisdom is reflected in today's methodologies such as [Schreiber *et al.*, 1999] that align technology to the business context, if it is needed to solve a knowledge-related issue at all.

In order to find out the position of ontology-based technology in the KM market, we extensively researched the KM literature both in print and on-line and visited in two consecutive years the largest European KM event (KM Europe) aimed at bringing together practitioners with vendors of software solutions. (For an exhaustive reference to KM information sources, projects and technologies, we also refer the reader to deliverables of the VISION strategic roadmap project [5].) While we found prototypes of ontology-based KM applications in the ontology literature, very few of the KM sources even mentioned the use of ontologies. Those which did, usually classified it under 'future KM technologies'. With this title, most sources suggested that ontologies have not made the step out of the lab yet and therefore no real cases should exist at present.

There is a danger for the ontology community in reading these statements as evidence to the fact that the KM market is not yet ready for making use of ontology technology. We would like to argue that reality is on the contrary: KM technology

---

[4] The KnowledgeStorm directory of business IT solutions lists over 3000 offerings under the keyword 'knowledge management'. Source: www.knowledgestorm.com

[5] The EU VISION project (EU IST-2002-38513). See `http://km.aifb.uni-karlsruhe.de/fzi/vision/`.

has independently reached a stage where there is a large potential for ontology-based methods.

3.1 The KM technology market

For a better insight into the current use of technology in the KM Market, we have selected 10 vendors from KM World's list of the top 100 companies in Knowledge Management, who are also regular exhibitors at KM World and represent a cross section of the entire market. We collected information on their technology use by sending out a questionnaire (see the Appendix) or in cases where we received no direct response, by looking through public information provided on company websites and in product leaflets. The summary of this survey is shown in Figure 1.

Two major stream of developments can be observed in KM technology: the static and dynamic management of knowledge. This reflects the dichotomy of knowledge we discussed in the previous section - static solutions support the knowledge as stock approach, while the dynamic stream emphasizes the flow nature of knowledge.

The static field of Knowledge Management is concerned with explicitly represented information, under the heading of (Enterprise) Content (or Document, Intellectual Capital) Management. This area gains popularity as companies are trying to find the way out of the 'infoglut': the growing mass of unstructured, unorganized information on the desktops of knowledge workers and in various systems of the modern enterprise. [6] Advances in this area are thus marked by ever more sophisticated models to structure, organize and find corporate or personal information. The state-of-the-art in knowledge representation is constituted by taxonomies (also called 'knowledge networks', 'knowledge modules' etc.), where a taxonomy is largely defined as a simple tree hierarchy of concepts. (Although some products feature a fixed set of cross-taxonomical relationships, for example to express that a certain concept -a node of the tree- is related to another concept.) Taxonomies are primarily used to classify documents and are often linked to vocabularies to support this process.

Taxonomy-based Knowledge Management therefore requires the acquisition of metadata to populate the taxonomy. This is carried out through a combination of Natural Language Processing (NLP), text analysis and machine learning. These are the key technology in classification tools from vendors such as Autonomy, Gammasite or Mohomine [Group, 2002]. The acquired metadata is stored in repositories in most cases along with the content. Metadata is stored internally in relational or object-relational databases and exposed mostly in XML through custom APIs, and more recently Web Services. Content repositories (such as the middleware platform of Documentum) also offer a number of management services to applications such as versioning, security, replication, workflow support etc.

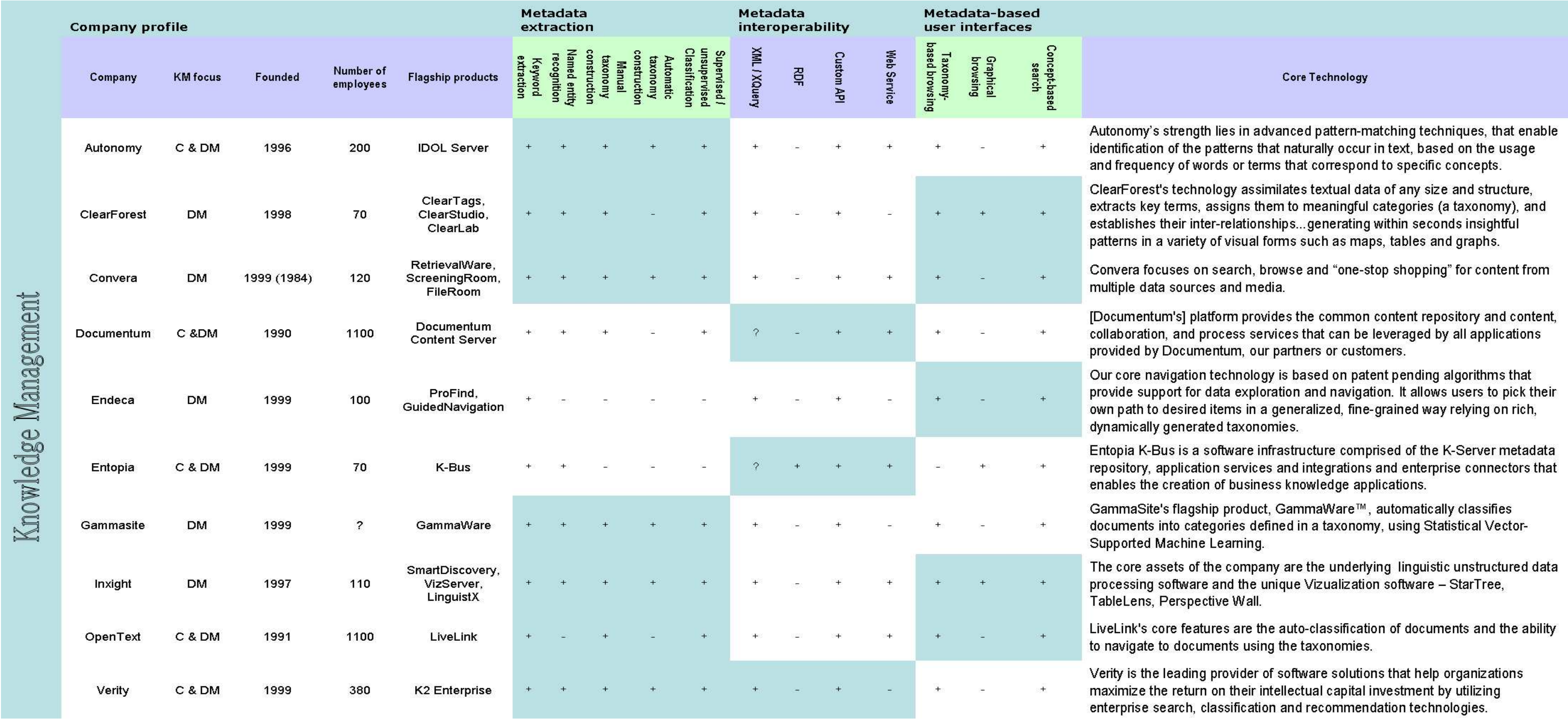
The metadata and content together form the basis of custom applications, where the core features are concept-based search and graphical browsing using the taxonomy. [7] Presentations may include creative graphical visualizations of large tree structures, guided navigation[8] and the use of highlights to draw attention to the entities found in

---

[6] The percentage of unstructured data within corporate information is an estimated 85% and growing. Source: The Delphi Group.

[7] Search engine vendors such as Verity, Factiva and Convera focus on metadata to improve on the precision of (Boolean) keyword-based search.

[8] Guided navigation refers to browsing interfaces that adapt the presentation based on previous selections



Company profile					Metadata extraction					Metadata interoperability				Metadata-based user interfaces			
Company	KM focus	Founded	Number of employees	Flagship products	Keyword extraction	Named entity recognition	Manual taxonomy construction	Automatic taxonomy construction	Supervised / unsupervised Classification	XML / XQuery	RDF	Custom API	Web Service	Taxonomy-based browsing	Graphical browsing	Concept-based search	Core Technology
Autonomy	C & DM	1996	200	IDOL Server	+	+	+	+	+	+	-	+	+	+	-	+	Autonomy's strength lies in advanced pattern-matching techniques, that enable identification of the patterns that naturally occur in text, based on the usage and frequency of words or terms that correspond to specific concepts.
ClearForest	DM	1998	70	ClearTags, ClearStudio, ClearLab	+	+	+	-	+	+	-	+	-	+	+	+	ClearForest's technology assimilates textual data of any size and structure, extracts key terms, assigns them to meaningful categories (a taxonomy), and establishes their inter-relationships...generating within seconds insightful patterns in a variety of visual forms such as maps, tables and graphs.
Convera	DM	1999 (1984)	120	RetrievalWare, ScreeningRoom, FileRoom	+	+	+	+	+	+	-	+	+	+	-	+	Convera focuses on search, browse and "one-stop shopping" for content from multiple data sources and media.
Documentum	C &DM	1990	1100	Documentum Content Server	+	+	+	-	+	?	-	+	+	+	-	+	[Documentum's] platform provides the common content repository and content, collaboration, and process services that can be leveraged by all applications provided by Documentum, our partners or customers.
Endeca	DM	1999	100	ProFind, GuidedNavigation	+	-	-	-	-	+	-	+	-	+	-	+	Our core navigation technology is based on patent pending algorithms that provide support for data exploration and navigation. It allows users to pick their own path to desired items in a generalized, fine-grained way relying on rich, dynamically generated taxonomies.
Entopia	C & DM	1999	70	K-Bus	+	+	-	-	-	?	+	+	+	-	+	+	Entopia K-Bus is a software infrastructure comprised of the K-Server metadata repository, application services and integrations and enterprise connectors that enables the creation of business knowledge applications.
Gammasite	DM	1999	?	GammaWare	+	+	+	+	+	+	-	+	-	+	-	+	GammaSite's flagship product, GammaWare™, automatically classifies documents into categories defined in a taxonomy, using Statistical Vector-Supported Machine Learning.
Inxight	DM	1997	110	SmartDiscovery, VizServer, LinguistX	+	+	+	+	+	+	-	+	+	+	+	+	The core assets of the company are the underlying linguistic unstructured data processing software and the unique Vizualization software – StarTree, TableLens, Perspective Wall.
OpenText	C & DM	1991	1100	LiveLink	+	-	+	-	+	+	-	+	+	+	-	+	LiveLink's core features are the auto-classification of documents and the ability to navigate to documents using the taxonomies.
Verity	C & DM	1999	380	K2 Enterprise	+	+	+	+	+	+	-	+	-	+	-	+	Verity is the leading provider of software solutions that help organizations maximize the return on their intellectual capital investment by utilizing enterprise search, classification and recommendation technologies.

Figure 1: Product portfolio of ten leading KM vendors in Document Management and Collaboration. Highlights show the technology focus of the companies.

the document text. Documents are almost always presented in the context of related concepts or documents.

In the dynamic area, the focus has been on action: connecting people to experts and expertize and facilitating direct exchanges. This is the target of collaborative applications (groupware), project management and workflow solutions, messaging services and more recently social network applications.

Looking at the KM market as a whole, we see the following picture emerging. Smaller companies in the static space contribute particular technologies to one of the three steps of the content management process (information extraction, taxonomy-based classification and visual information access). Larger vendors tie these technologies together in comprehensive solutions and combine them with collaborative functionality. This has the added value of integrating community activity with the more static sources of knowledge. In practice, this means that information about documents, people, workflows, projects etc. all become metadata interlinked with one other.

It is particularly revealing to compare this technology portfolio with the generic architecture of ontology-based systems (shown in Figure 2) that we have seen recurring in the past both in the literature (e.g. [Abecker and van Elst, 2003] and in the developments of a number of European research projects that we have been involved in[9] From this comparison, it is immediately apparent that ontology-based systems mirror the architecture of KM software (and vice versa), with the addition of explicit representations of metadata, which also allows the use of generic reasoning engines. Knowledge Management systems could thus could thus benefit from the standardization work that focuses on the interoperability between system layers. This includes standard languages for ontology representation such as the RDF and OWL languages recommended by the W3C and work on the standardization of query and rule languages for the Semantic Web. What is lacking on the ontology side is the experience in the building of ontology-based systems that would allow for robust methodologies of creating and maintaining ontologies and ontology-based systems in an organization context. (Up to now efforts at creating methodologies were typically tied to a particular set of tools that emerged as a result of research efforts. This will hopefully change with the chartering of a working group within the W3C to collect best practices that can be later aggregated and abstracted.)

Looking more closely at the technology used by KM vendors, two key concepts meet the eye that can be used as starting points for introducing and arguing for ontology technology: XML and Taxonomies. In the following section, we examine both options in more detail.

3.2 XML vs. RDF/OWL

XML is a well-understood and widely acknowledged technology in the entire domain of enterprize information systems. Software engineers appreciate XML for the variety of complementary tools and technologies such as XML databases, standard schema and query languages, a range of editors, translators, processors etc. Information Officers value XML for reducing the cost of interoperability between systems and services.

The high status of XML compelled also the W3C to adopt it also as a notation for

---

of the user. See for example Endeca or Aduna (`http://www.aduna.biz`).

[9] For example, On-To-Knowledge, SWAP and SEKT.

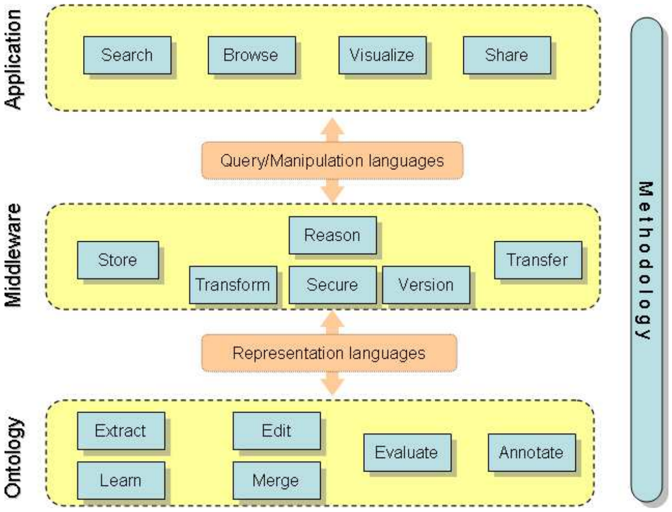


Figure 2: A generic architecture of ontology-based systems.

ontologies. This kind of contingency is popularized by the famous Semantic Web layer cake diagram in Figure 3[10]. This picture shows the vision of the Semantic WEb as a set of languages building upon existing standards such as XML, URIs and Unicode. Despite the good intentions, however, this diagram hides completely the true relationship between XML and ontology languages from the W3C (such as RDF and OWL). By doing so it has done more damage to the argument for ontology languages over XML than any other conceptualization of the next-generation Web.

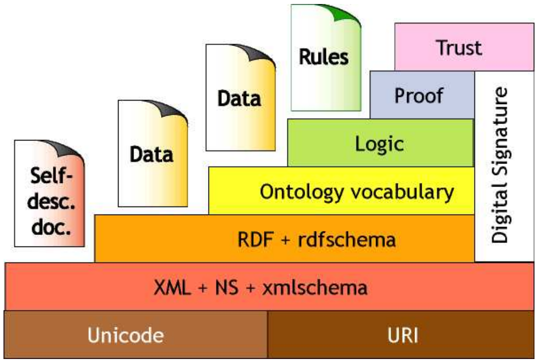


Figure 3: The Semantic Web layer cake.

The XML data model is a (directed) tree defined by the nesting of elements, with the added possibility of links across branches using the relatively cumbersome ID/IDREF mechanism. This model originates from the predecessor of XML called SGML, which was mainly used for marking up large documents. Text documents follow the tree structure themselves as paragraphs are nested in subsections, subsections are nested in sections, sections are nested chapters etc.

---

[10] The diagram has appeared in presentations of the W3C such as `http://www.w3.org/2001/09/06-ecdl/slide17-0.html`.

On the other hand RDF, which is also the basis of the more expressive OWL language, proposes a data model based on arbitrary graphs built from single edges between the nodes representing classes or instances. This structure is better suited for creating conceptual domain models, which are often so rich in cross-taxonomical relationships that a hierarchical representation would prove to be difficult. Schema languages for RDF also transcend the relatively vague notions of XML and its schema languages (such as entities, attributes, nesting etc.). Instead, RDF(S) and OWL provide a more suitable set of constructs (classes, instances, properties, subclasses), which also have a precisely defined interpretation.

RDF(S) is thus rather an advancement than an extension of XML in terms of knowledge representation capabilities. In fact, RDF(S) relies on XML merely as a notation: in the XML-based serialization of RDF graphs the edges (or short paths) from the graph are simply written out as a flat list. [11] This means that there is an N:1 correspondence between XML serializations and RDF graph models. Moreover, the XML tree reflects only a particular representation and has no specific relation to the semantics of the ontology. This also means that no generic schema can be given for RDF/XML, as such documents mix elements of the notation with elements of the domain (which is an infinite set). While RDF/XML documents can be stored in XML databases and queried using XML Query languages such as XQuery or transformed using XSLT, these queries and transformations will be sensitive to a particular serialization. All in all, the only useful tools with respect to RDF/XML files are XML editors and they are only helpful in checking the well-formedness of the representation.

While the transition from XML to RDF-based representations of taxonomies is a potentially larger and more extensive step then the SW layer cake suggests, it has considerable benefits to KM products, both in integrating systems and information. A standard representation of knowledge models and instance information (metadata) contributes to better interoperability between taxonomy products from different vendors, which will become more important as the taxonomy market consolidates and standard solutions emergence. On the short term, the main push will come from the need to integrate more easily the metadata created externally, often by the application that produces the content itself. An example of such an application is the latest version 6.0 of Adobe Acrobat, which attaches metadata to documents in RDF format using the Dublin Core ontology [Dublin Core, 1999]. The ability to reuse such simple metadata and to integrate it with other kinds of metadata available in content and collaboration systems (such as knowledge about the content, author, project, workflow etc.) will together drive the need to integrate RDF capabilities into existing KM solutions. As we will see in Section 5.2, integration is also the basis for reasoning: revealing patterns or insights in the knowledge that is not possible by looking at the single sources of such knowledge.

3.3 Taxonomies vs. Ontologies

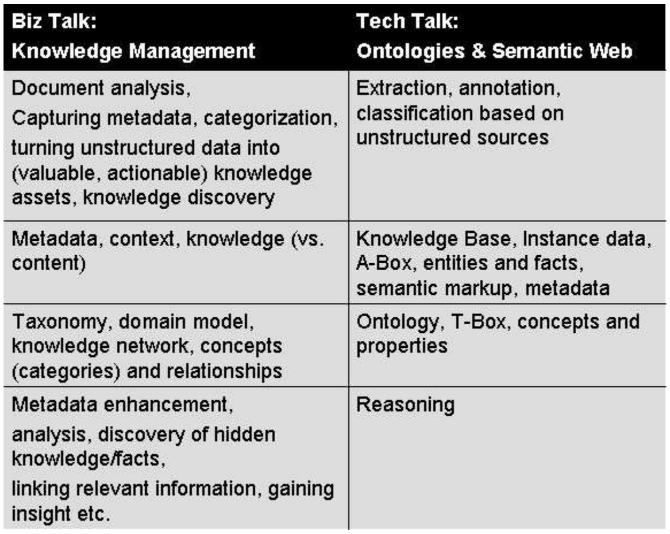
The second argument examines ontologies as an advancement in conceptual modelling over taxonomic structures. As we have demonstrated, taxonomies are well-established tools of KM experts supported by a number of applications.

Taxonomies lend as a natural starting point in explaining ontologies as they can

---

[11] There are other notations of RDF such as N3 or N-Triples, which are not based on XML. These are often preferred over RDF/XML for easier authoring.

be easily conceived as a sort of lightweight ontology. In fact, a look at the vocabulary used in the business and scientific literature reveals that much of the great divide between the two communities is in reality a difference in vocabulary (see Figure 4). Once we translate the business terminology, the differences between an ontology and a taxonomy boil down to a more principled design of conceptual models.



Biz Talk: Knowledge Management	Tech Talk: Ontologies & Semantic Web
Document analysis, Capturing metadata, categorization, turning unstructured data into (valuable, actionable) knowledge assets, knowledge discovery	Extraction, annotation, classification based on unstructured sources
Metadata, context, knowledge (vs. content)	Knowledge Base, Instance data, A-Box, entities and facts, semantic markup, metadata
Taxonomy, domain model, knowledge network, concepts (categories) and relationships	Ontology, T-Box, concepts and properties
Metadata enhancement, analysis, discovery of hidden knowledge/facts, linking relevant information, gaining insight etc.	Reasoning

Figure 4: A comparison of the terminology of Knowledge Management and ontology research.

At this point, one could take an advantage of the fact that the term ontology is mostly unknown in the KM world and argue for the introduction of the ontologies as conceptual models that improve on the the weak, unfounded taxonomic structures known to KM.[12]. Technically, this can also be avoided by omitting the word ‘ontology’ entirely and discussing principled, standardized taxonomies that use a richer set of constructs with well-defined semantics. A ‘good’ and a ‘bad’ taxonomy can be easily explained to a KM audience by providing examples. See Figure 5 for an example.

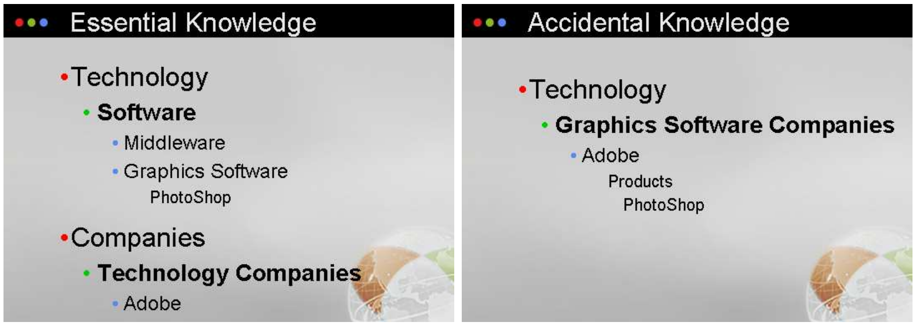
Further, we contend that knowledge engineers are largely aware of the fact that the core of a good taxonomy should reflect a wide consensus over the domain, since such taxonomies are likely to be more stable. This means that systems built around such models are easier to maintain and more likely to be accepted by the community of users. Personal views (contexts) can be easily derived from or linked to such taxonomies.

Lastly, in promoting ontologies one may also adopt the strategy of shifting the focus from the nature of the technology and opt for technology-neutral names reflecting application scenarios instead. This tactic is particularly appropriate in the early days of technology push when interoperability using standards is a low priority.[13] This is also the preferred approach in promotion towards a non-technical audience that is more interested in what the technology can achieve rather than knowing its particular nature.

---

[12] This is the approach chosen by SW start-up Semagix, who write “The Ontology is a superset of a Taxonomy...”)

[13] As a representative of Network Inference, a UK-based Semantic Web startup, explained: “As soon as the day comes when someone would call us asking for OWL or the Semantic Web, we will include it in our brochures.”



Essential Knowledge

- Technology
  - **Software**
    - Middleware
    - Graphics Software
      PhotoShop
- Companies
  - **Technology Companies**
    - Adobe

Accidental Knowledge

- Technology
  - **Graphics Software Companies**
    - Adobe
      Products
        PhotoShop

Figure 5: A good taxonomy is a stable hierarchy of the basic components of the world, organized by their essential natures. Courtesy of Joshua Powers, Chief Ontologist at Convera.

Part of the difficulty of arguing for ontologies under any name is that ontologies are deeply buried behind the user interface and even there form a rather static component of a system. Naturally, this position attracts questions from managers such as “What does this *do* to help us manage the knowledge we need?”. This means that arguments for ontologies will need to be phrased in active terms, that is, by looking at the knowledge processes and scenarios supported by this technology. This is the topic of the following section.

4 Ontologies in action

What is the innovative value of ontology-based methods, compared to the existing and well-understood tools of KM ranging from whiteboards to expert systems and content classification solutions? As noted before, making a compelling argument for ontologies in face of this question requires finding those knowledge processes and scenarios that can directly benefit from ontologies.

In order to crystallize the potential of ontology-based methods, we first need to take a closer look at the fundamentals of ontologies. This will help us to identify the role and the scope of ontology use in KM applications and will lead us to a simple classification of applications of ontologies, which we will call the Semantic Web Matrix.

An ontology, by its most cited definition in AI, is a shared, formal conceptualization of a domain [Gruber, 1993; Borst *et al.*, 1997]. Ontologies are data models with two special characteristics, which lead to the notion of shared meaning or semantics:

1. Ontologies build upon a shared understanding within a *community*. This understanding represents an agreement of experts over the concepts and relationships that are present in a domain. (The human factor in ontology-based KM.)

2. Ontologies use machine-processable representations (expressed in formal languages such as RDF [Lassila and Swick, 1999] and OWL [Dean *et al.*, 2004]), which allows computers to manipulate ontologies. This includes transferring ontologies among computers, storing ontologies, checking the consistency of ontologies, reasoning with or with the help of ontologies etc. (The machine factor in ontology-based KM.)

These factors make it possible to create the semblance of intelligence in applications by incorporating domain knowledge in the form of an ontology. Note that the two factors carry different weights in this respect. While machines can manipulate and reason about domain knowledge based on the ontology of the domain, knowing what the symbols and rules stand for remains a human function. Therefore no ontology may exist without a community supporting it.

The first item above also provides the starting point for our investigation into the functionality supported by ontologies, which makes up the first dimension of our classification. We propose that a shared understanding represented by ontologies can be used for driving three key processes of Knowledge Management: Communication, Integration and Reasoning (see Figure 6). As we will discuss later, these processes build on each other as layers of a pyramid and represent increasing levels of formality in terms of the ontologies involved as the complexity of the knowledge required increases. In this respect, this dimension is related to the different levels of formality of ontologies, ranging from simple vocabularies to full-fledged logical structures allowing powerful inferences [Smith and Welty, 2001]. The semantic dimension will form the vertical axis of our matrix as described in the following section.

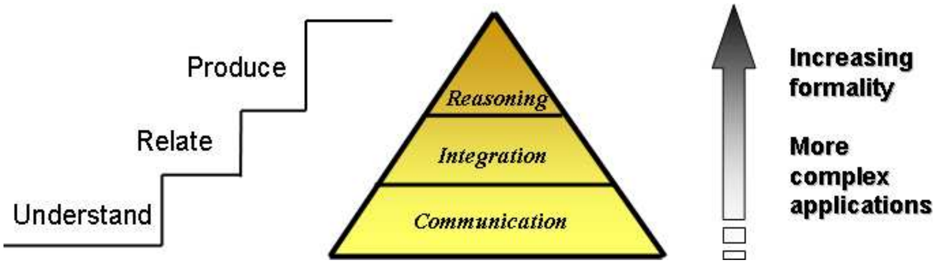


Figure 6: Ontology functionality in applications ranges from ontology usage for Communication to the most complex cases of Reasoning.

The second point above tells us about the application scope of ontologies. On the negative side, the application of ontologies is limited to explicit knowledge. Nevertheless, once we have some kind of knowledge described in an explicit, formal way, we can trust our computers to work with this knowledge. This not only means an increase in the size of the knowledge we can handle, but more importantly gives us the opportunity to use our computer networks (intranets, extranets and the internet) to significantly extend the reach of this knowledge. Briefly put, this means communication across autonomous groups, integration of their knowledge and reasoning by aggregating information from disparate sources. This development (dubbed the Semantic Web) has several stages that we describe in more detail in Section 6. This second, web dimension will add the horizontal axis to our Semantic Web Matrix.

Our framework assimilates key aspects of Uschold's earlier work with respect to the intended purpose or benefits of using an ontology and the various of levels formality required [Uschold and Jasper, 1999]. We do not differentiate, however, by the orthogonal issues of representation and supporting technologies. Further related work include the excellent analysis of Abecker and Elst on ontologies for Knowledge Management [Abecker and van Elst, 2003]. While we fully concur with their analysis, we extend their work with an investigation of decentralized ontology-based systems, the above mentioned Web dimension.

5 The Semantic dimension

In the following we go through the three key knowledge processes that ontologies support in applications. These processes are visually laid out as a pyramid as they are directly dependent on underlying processes (see Figure 6. As a result, the complexity of the applications increases –and the number of real cases decreases– towards the top of the pyramid.

5.1 Communication

Ontologies represent a shared understanding among members of a community about a set of terms and the descriptions created with those terms.[14] Thus ontologies potentially provide the means of transportation in the process of communicating knowledge across the domain.

Ontologies facilitate communication by providing shared notions that can be used to formulate statements (queries, replies etc.) about instances in the domain. For the receiving party of such a communication, the ontology helps to understand messages and to avoid ambiguities by providing the interpretation context. In general, by leveraging an accepted, well-understood language organizations can learn better, become more aware and respond faster to their environment.

Ontologies used for communication need not be very expressive. In fact, lightweight representations such as topic hierarchies or Mind Maps [Buzan and Buzan, 1996] are often preferred over intricate formalisms since they can be authored by the domain experts themselves. The direct involvement of the experts in the creation of an ontology and the ensuing discussions help identifying differences in understanding, regardless whether the ontology is used to describe a domain, a business process or an entire organization. Through this process communities can ensure that members of a group, department or the entire organization commit to the same shared view of the domain, task or enterprize. Moreover, through the building of the ontology a part of tacit knowledge becomes explicit in a natural, nonobtrusive way.

When applied consistently, a shared and formally acknowledged language enables working groups (e.g. a first- and second-line helpdesk) to cooperate more efficiently by avoiding misunderstandings. The method can be extended in scope to inter-departmental communication or messaging with partners. One only needs to be careful not to repress creativity and experimentation by an overly elaborate, broad and rigid ontology.

This application area of ontologies draws less attention from the Knowledge Representation and Reasoning community, since the low level of formality allows to make

---

[14] In the following we will refer to the term ontology as the entire structure of classes, instances and their interrelationships and make no distinction among the schema level (the level of classes) and instance data. In some representation languages and systems the description of classes and their relationships is kept separate from instance information and in that case the term ontology is used for speaking about the level of concepts only. The term metadata (data about data) is used for the instance data in the specific case where the ontology is used to describe and normalize a set of instances (e.g. documents).

only trivial inferences. Nevertheless, software applications can assist in authoring the ontology in a machine processable format so that it can be archived, transferred and retrieved later. Machine representation also has the advantage that the descriptions which use the ontological terms can be checked for correctness against the ontology.

Example 1: In the work of Gordijn (see [Gordijn and Akkermans, 2001; Gordijn and Akkermans, 2003]), the e^3*value* ontology is used to describe innovative e-commerce ideas in the exploration phase. Representing the business proposal in terms of a standard ontology ensures a shared understanding by the various stakeholders involved and thus improves decision-making. Formal representation and the machine processing of such business models is an advantage for handling large, complex cases and it enables quick checking of certain modelling assumptions. However, automation is not a strict necessity for this use case. Although business modelling tools have been developed and appear to offer welcome assistance, the shared process of communication, understanding and subsequent decision making is the really crucial element.

Example 2: As an industrial partner in the On-To-Knowledge European research project SwissLife, a Swiss insurance company created an ontology-based system for the retrieval of passages from the International Accounting Standard (IAS) [Reimer *et al.*, 2003]. The IAS contains regulations used by accountants all over Europe. However, only the regulations relevant to Swiss Life made up more than 1000 pages of text. Besides the size of the corpus, the problem was that the document uses a very specific vocabulary and accountants across the global firm had problems choosing the right terms when searching inside the standard.

For the operation of the firm it was crucial that employees could find in a quick and reliable way the relevant text passages within the document that explain how the IAS rules are to be applied in a particular situation. The ontology here helped to improve the communication between (authors of) the text and the accountants, who could rely on the ontology in phrasing their queries to the system.

5.2 Integration

Ontologies are more than a simple vocabulary of precisely defined notions: their true power is vested in the description of relationships between entities of the domain. As the inventor of the World Wide Web writes in [Berners-Lee *et al.*, 1999], “In an extreme view, the world can be seen as only connections, nothing else. We think of a dictionary as the repository of meaning, but it defines words only in terms of other words. I liked the idea that a piece of information is really defined only by what it’s related to, and how it’s related. There really is little else to meaning. The structure is everything.”

Full-fledged ontologies rich in relations are more powerful than the simple taxonomic hierarchies that appear in current KM solutions. Relations (links between concepts) increase tremendously the number of ways to navigate and search through the domain, analyze, classify and visualize knowledge.

To harness this rich web of concepts for information integration, heterogeneous resources need to be marked up by the terms of the ontology. This process, also called *semantic normalization*, clusters conceptually related pieces of information, regardless of formats and representations. While the effort is similar to attaching keywords or classification codes to information, the resulting descriptions have precise interpretation vested in the ontology.

The term ‘resource’ is used here in the broadest possible sense, including databases, textual content, user profiles, web services etc. In enterprise KM applications, the most promising developments are in the area of Organizational Memories (OM)[15], since these systems typically need to deal with heterogeneous and unstructured sources of knowledge, including documents, emails, web pages, calendars, meeting notes etc. Heterogeneity makes it difficult to discover interrelationships (“knowing what we know”), while unstructured formats (designed for humans to read) are not machine searchable. In fact, the technology is particularly appealing for the description of multimedia such as images and video fragments, whose subject matter is directly not accessible to machines (see e.g. [Schreiber *et al.*, 2002])

Attaching ontology-based descriptions (metadata) to existing information may at first seem to contribute to the very information overload it is trying to solve. The richer ontologies used in this scenario also represent more difficulty in terms of construction with respect to the previous case. The ontology has to be elaborate enough to cover all sources to be used. Our practical experience with ontology building also shows that determining, naming and scoping relationships are very challenging modelling tasks.

On the positive side, systems benefit from more precise, concept-based searches and browsing. The domain ontology used to classify information can be used to create intelligent interfaces that help the user in formulating queries in ontological terms [Iosif *et al.*, 2003]. In effect, the end user notices only a more intelligent interface that is aware of the relations between concepts in the domain, without having to be knowledgeable of the technology behind. Indeed, even an expert who is only partially familiar with the domain may learn the meaning of concepts by interpreting them in relation to other known concepts. The benefits may thus outweigh the overhead of maintaining rich ontologies and metadata.

We place the integration scenario on top of the communication layer, since integration relies on a mediated communication process. In this case, the communication occurs between the creator of the ontology and the descriptions (metadata), and the actor who performs the retrieval using the same ontology. In most cases, this communication is mediated by a computer system that helps the user in forming his query, by suggesting alternative terms, generalizations, specializations etc. based on the machine’s understanding of the ontology language. The improved access to the organization’s collective knowledge results in higher knowledge productivity of workers as they (re)use existing knowledge more effectively and defend themselves from information overload.

Example 1: EnerSearch, a Sweden-based virtual organization researching issues related to the use of IT in the energy utility segment, created a semantic portal as a case study in the European On-To-Knowledge project [Iosif and Mika, 2002; Iosif *et al.*, 2003]. By its charter, EnerSearch is responsible for the dissemination of its research results (contained in over a hundred publications) to its shareholders, its members and the general public, who all access this content via the EnerSearch web.

However, keyword-based search and fixed browsing proved to be a bottleneck in tapping into this content. Given the value of the knowledge contained therein and its specific nature (scientific research in a well-defined domain), EnerSearch decided to invest in ontology-based technologies to improve the effectiveness of both browsing

---

[15] Some authors, such as Dieng [Dieng *et al.*, 1999] consider Skills Management applications as the subset of OM applications, where the primary entities are people instead of documents.

and search. A semantic portal and query interface were built on a model that integrated metadata which was already available in the organization with an ontology automatically extracted from the corpus. An evaluation of the system showed a gain in user satisfaction through the use of the new facilities.

Example 2: As part of their role in the DAML Program[16], researchers of the University of Maryland Baltimore County have developed ITTALKS [Cost *et al.*, 2002]; a web portal that offers access to information about information technology (IT) related events. ITTALKS provides users with numerous details describing the IT events, including location, speaker, hosting organization, and talk topic. The portal also stores user profiles in order to provide personalized matches to its visitors.

The information contained in the site comes either from the users of the site or as submissions from agents crawling various websites with related information. ITTALKS uses a number of simple ontologies from the DAML collection to represent this information, including the ACM Computing Classification System[17], which is used to classify events. The system is accessible for semantics-based search and navigation for both humans and machine agents, through the portal and an agent system, respectively.

5.3 Reasoning

Applications of terminological reasoning represent the most complex use case for ontologies. While the lower two layers in the pyramid build on knowledge about *what* the kinds of objects are in our domain (communication) and *how* they're related (integration), this scenario requires knowledge on *why* they are related. In other words, this scenario concerns the rules and principles behind a certain conceptualization.

In the KM literature, the difference is concisely articulated by Russell Ackoff, who –unlike many of his contemporaries– classifies the human mind in five states: data, information, knowledge, understanding and wisdom. The difference between what Ackoff calls 'knowledge' and 'understanding' is subtle, but important. As he writes about knowledge in [Ackoff, 1999],

"Knowledge is the appropriate collection of information, such that it's intent is to be useful. Knowledge is a deterministic process. When someone 'memorizes' information (as less-aspiring test-bound students often do), then they have amassed knowledge. This knowledge has useful meaning to them, but it does not provide for, in and of itself, an integration such as would infer further knowledge. For example, elementary school children memorize, or amass knowledge of, the 'times table'. They can tell you that '2 x 2 = 4' because they have amassed that knowledge (it being included in the times table). But when asked what is '1267 x 300', they can not respond correctly because that entry is not in their times table. To correctly answer such a question requires a true cognitive and analytical ability that is only encompassed in the next level..."

The difference between knowledge and understanding is thus analogous to that between memorizing and learning. Knowledge is information that is relevant (in context) and passes all other customary filters we apply when deciding whether to incorporate information into our knowledge base [Godbout, 1999]. Understanding, on the other hand, is analytical and interpolative. It concerns the capacity to synthesize new

---

[16] The US DAML project. See `http://www.daml.org`

[17] http://www.acm.org/class/

knowledge from what is previously known, a symptom often associated with the term 'intelligence'.

The understanding of a domain is very difficult to explicate and requires the highest skill from the knowledge engineer in elicitation. Whenever rules and principles (a.k.a 'business logic' or 'business model') are built in line-of-business applications, they are typically buried inside the code in a way that requires significant re-engineering efforts to recover.

Deeper understanding of a domain leads to more compact models; confer storing the times table against storing the procedure for multiplication. Embedding this understanding in the form highly axiomatized, “rich” ontologies adds the value of easier maintenance of applications: whenever the underlying conceptualization of the domain changes only the ontology needs to be evolved. The code of the application, however, remains unaffected.

Analytic applications of rich ontologies include the checking of the consistency (conformance) of data against the model. This is the basis of applications suggested for stock trading authorities to analyze filings and stock exchange data to alert for insider trading [Ewalt, 2002] or created for law enforcement agencies for the threat assessment of flight passengers [sem,].

Synthetic type of applications, such as classification (which is directly supported by Description Logics (DL), among them the new standard ontology language OWL DL [Dean *et al.*, 2004]) allow to automatically find the correct place for a concept (such as a product description) within a hierarchy of descriptions. This service may be used during ontology development e.g. to automatically place newly created complex concepts, as well as in data entry or in answering complex queries formulated as partially specified concepts. Clearly, the usefulness of such reasoning depends on the richness of the specification of the new concept.

Highly formal, rich ontologies have first made headway in fields of science such as bioinformatics, where the potential gains in linking knowledge offset the costs of elicitation and formalization. We classify applications belonging to this scenario by a need for formal knowledge to reason about the domain. Note that ontology-based reasoning shows its own potential compared to traditional expert systems when used to analyze or synthesize knowledge from diverse sources, possibly supplied by different communities. This is also the case in the examples below.

Example 1:

In the TAMBIS project researchers of the University of Manchester implemented a system that provides transparent information retrieval and filtering from a number of information sources used in bioinformatics. The system interacts with the user in terms of a generic ontology of molecular biology, also developed within the project. Queries in terms of this ontology are translated to particular requests against the various information sources. The system creates the illusion of a uniform service by integrating heterogeneous sources and performing complex reasoning within the domain.

The ontology created in the project uses the highly expressive GRAIL ontology language. The expressiveness of GRAIL means that besides a basic asserted hierarchy most concepts can be constructed in some way from existing concepts. Such composed concepts are automatically checked, classified and maintained by the GRAIL reasoner. Powerful constraints called ‘sanctions’ in GRAIL even make it meaningful to generate or infer concepts based on existing definitions.

Example 2:

Automating the complex tasks of software discovery, monitoring, configuration, matching and composition is a grand challenge for both Enterprise Application Integration (EAI) and the Semantic Web. A number of technical architectures have been proposed to meet this challenge, based on the notions of Web Services, the Grid or Multi-Agent Systems (MAS).[18] However, common to all approaches is the need for the formal descriptions of functional and non-functional characteristics of software components. (Non-functional characteristics include issues related to security, trust, value, cost, ratings, availability etc.)

The complexity of describing software components and their potential interaction means that these descriptions require high expressivity from the underlying knowledge representation formalism, especially in face of the variety of tasks to support [Peter Mika and Daniel Oberle and Aldo Gangemi and Marta Sabou, 2004]. Formal descriptions of task knowledge and method (component) functionality have been considered earlier under the heading of Problem Solving Methods (PSMs) with similar calls for expressive formalisms. (See [Motta *et al.*, 2003] for a revisit of this work in the Semantic Web context.) Additionally, software descriptions on the Semantic Web rely on rich domain ontologies to precisely characterize their input/output parameters and the preconditions and effects of their execution, if semantic compatibility is to be achieved in the large, open domain of the Web.

6 The Web dimension

By giving information unprecedented availability and a friendly interface, the World Wide Web has drastically changed the way we go about solving information-intensive tasks in our lives. In this pretext, the vision of the Semantic Web is often explained as an evolutionary step in web technology: an extension to the current “syntactic” web with a network of ontologies crawled by intelligent agents acting on behalf of their users [Berners-Lee *et al.*, 2001].

In this paper we would like to promote a complementary view of this development. In accordance with our application-oriented analysis provided in the previous section, we describe this advance as a technical development towards distributed knowledge-based architectures, but rather as a decentralization of our ontology-based scenarios. We believe that such a definition would prove more useful in arguing for Semantic Web technology in all application areas, of which the World Wide Web is one special case. In particular, companies may draw significant benefits from semantic technology in certain domains without relying on an implementation of the grand vision of the Semantic Web. We choose to complement our classification framework with the Web dimension so that we can identify the advantages and disadvantages of moving a KM application towards more and more decentralized settings. We describe the various stages of decentralization in ontology-based applications, which mark the Web dimension of our framework. When our ontology functionality pyramid is given depth by this dimension, we reach the Semantic Web Matrix, a comprehensive classification of ontology-based application scenarios.

---

[18] For collective representation, see the W3C Web Services Activity `http://www.w3c.org/2002/ws/`, the Globus Alliance `http://www.globus.org/` and the Foundation for Intelligent Physical Agents `http://www.fipa.org/`, respectively.

We will show that a combination of decentralization and semantic technology may be argued from a theoretical point of view and it also promises benefits to the scenarios identified. However, managerial and technical challenges to this approach are also surfacing. We will argue in our conclusion that some of these challenges are here to stay, due to a fundamental contradiction between the large-scale distribution and strict formalization of knowledge.

6.1 Stages of decentralization

Decentralization means a dispersal of components across lines of authority. As Tim Berners-Lee, one of the original creators of the web writes in [Berners-Lee *et al.*, 1999]: "As with the current Web, decentralization is the underlying design principle that will give the Semantic Web its ability to become more than the sum of its parts."

This process has several extents in the case of ontology-based applications. In our analysis we identified four distinct stages of decentralization, which can be described as follows.

1. **Single authority applications** These are ontology-based applications that operate in a closed environment where resources (knowledge assets such as content, services etc.) and application logic are all under the control of a single entity.

   Note that this definition does not refer to the technical implementation of the application which might just as well take the form of a distributed system. An example of such a single authority application would be a peer-to-peer content network within a global organization.[19]

2. **Shared ontology, local resources** Applications that mark this stage build on a central ontology, but operate on local resources under independent control. Note that this scenario presupposes the existence of some form of a network that connects these resources to the application.

   We also consider here those applications that allow different views to be defined over a shared ontology or have multiple local ontologies that are aligned to a shared ontology. In these cases, the difference from the next stage, namely having independent local ontologies, is that the mappings between the ontologies are predetermined by their relationships to the shared mediating ontology.

3. **Local ontologies, local resources** Applications at this level of decentralization build on local resources as well as local theories (ontologies) over those resources. There is no global or shared ontology to which all ontologies could be mapped in a straightforward manner as in the previous case.

4. **Distributed application logic** These applications are characterized by the decentralization of the application logic itself. Nodes of the network not only take over control of resources and ontologies, but create their own applications from independent components (static services and dynamic agents) distributed over the network. This stage, where both data and logic are decentralized, represents the most significant loss of control over an application.

---

[19] See the NXT3 from Nextpage for a commercial implementation, `http://www.nextpage.com`

While distinguishing different stages of decentralization, we mention that decentralization can take place at several levels, i.e. on the individual level (social networking), group or community level, organizational level or global level. Thus there might exist scenarios where multiple stages of decentralization exist at the different levels. For example, one can imagine an application where members of a group share a common ontology while on the group level mappings take place with or without a mediator.

The stage of decentralization determines the place of an application along the Web dimension of our framework. When the previously described semantic aspects of applications is complemented with the Web dimension, we arrive at the Semantic Web Matrix as shown in Figure 7. The value of this matrix lies in its ability to characterize ontology-based systems along dimensions that have direct relevance for the Knowledge Management application scenario.

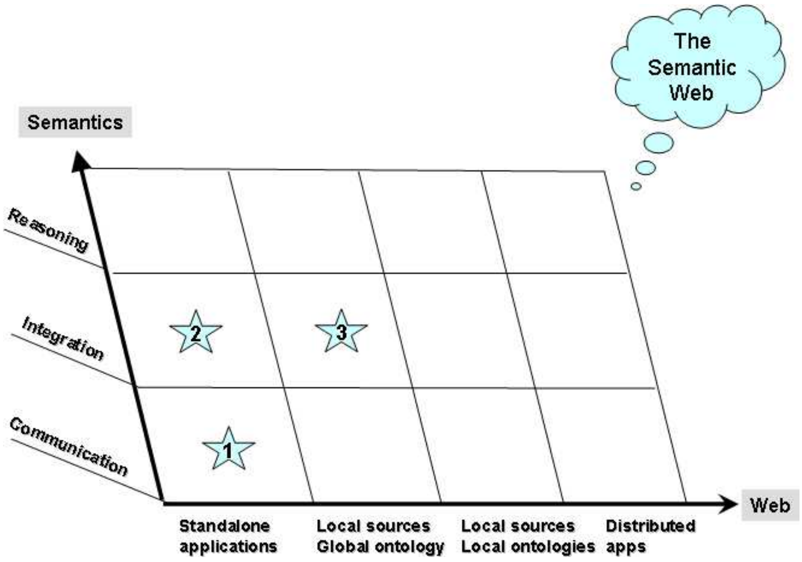


Figure 7: The Semantic Web Matrix classification framework for ontology-based KM applications. We marked the places of the SwissLife and e^3 applications (1), along with the EnerSearch (2) and ITTALKS (3) cases. We also showed the extent of activities of some major European ontology research projects.

Placing most of our examples in this framework is straightforward, with the exception of the IBROW project. In this case the ambition of the project for a distributed scenario is not reflected by the case studies completed, which both rely on single libraries of problem-solving methods designed for particular tasks. This meant that the examples could elude the challenge of mapping descriptions to UPML and matchmaking between tasks and methods.

Note that from a technical point of view, the task of the IBROW project was a special case of the general effort toward the semantic normalization, automated retrieval and composition of Web Services. The architecture and information flow of PSMs is fixed and the places where components need to be fitted are clearly marked, which is an easier task then the construction of arbitrary applications from Web Services.[20]

---

[20] A fixed design means that components can be explicitly described as modules that fit a certain hole in

The Semantic Web Matrix also allows to compare alternative notions of the Semantic Web as an application of distributed ontologies that reference existing web content.[21] Considering the most quoted vision of the Semantic Web [Berners-Lee *et al.*, 2001], we would place this development in the upper-right corner of the Matrix, as the scenario described in the article it requires distributed application components (agents, Web and Grid Services) working with highly expressive knowledge. By following the definition of the Semantic Web Challenge[22], we could relax this definition to include all ontology-based systems that meet or exceed the second stage of decentralization. (Web-based ontology applications that aggregate information from at least two, independently controlled information spaces.)

In the following section, we describe some of the advantages and disadvantages in bringing KM systems forward along the Web dimension.

6.2 Decentralized scenarios for ontology-based KM

There are practical as well as fundamental motivations for committing to a distributed style of knowledge management. From the practical side, we observe a number of trends which lead to a distributed accumulation of information and a loss of the control that would be necessary for centralized management. We just mention three of these trends as examples:

- **Pervasive computing.** Pervasive or ubiquitous computing refers to the growing number of embedded or mobile devices ('gadgets') that we integrate into our environment for more convenient access to computing resources. An embedded system is some combination of computer hardware and software, either fixed in capability or programmable, built into a particular device to execute a specific task, e.g. monitoring or control. A mobile device is the collective name for the number of portable devices (PDAs, cell phones, wearable computers etc.) used for personal information management and online connectivity on the move.

  While these devices are increasingly network-enabled, due to their sheer number it's hardly feasible to consolidate the information generated by them in a central repository. But even if it were technically feasible, it may not be sensible to do so. Mobile devices, for example, are known to be personal and trusted systems; users of these devices are quite likely to feel resentment over any centralization attempt. [23]

- **Virtual communities.** With the opening of markets, today's businesses compete in an international environment. Often the globalization of business activities

---

the architecture of a given method. The predefined role of components within methods implicitly provides certain relationships between them and even makes it possible to determine *a priori* certain dependencies between components and the behavior of the entire system, e.g. that one component results in larger result sets then the other.

[21] While the Semantic Web was originally conceived as an extension of the current Web infrastructures, it is now common to call ontology-based software 'a Semantic Web application' even if it does not have the goal to work on a global scale (e.g. intranet applications) and do not integrate existing web content. However, we wish to use the term here in the original sense.

[22] The Semantic Web Challenge. See `http://challenge.semanticweb.org`

[23] One might recall the rise and fall of the idea of the Network Computer (NC), thin clients promoted by vendors such as Oracle in the mid-1990s to lower the TCO of computer systems. Introductions of NC systems resulted in failure as users refused to relinquish their trusted, self-controlled desktop computers in exchange for centrally managed 'dumb' terminals.

resulted in geographically distributed centers of competence, which exchange knowledge through international projects and networking rather than by situated learning. These virtual communities thus resemble the distributed versions of the traditional Community of Practice described in Section 2.1.2.

Virtual communities are not only formed at a team level or for executing a particular task or project. Virtual communities of knowledge workers may also be formed outside organizational lines, e.g. for professional knowledge sharing. A Virtual Organization is a consortium of members that are geographically apart while appearing to others to be a single, unified organization with a real physical location.

Common to these communities is a flat, distributed structure in which resources and knowledge are under the control of the individual members of the community. Furthermore, they coordinate their efforts not as a result of control, but as a result of their own motivation to increase the effectiveness of their operations, which results in a volatile structure. The role of information technologies in supporting virtual communities is thus enabling interaction, instead of centralizing knowledge through portals or repositories.[24]

- **The Web legacy.** The web as we know it today is a giant distributed document store, with files dispersed on an estimated 171 million hosts[25]. A large part of the success of the web is attributed to the idea of local control. Servers on the web are all independently managed by their respective owners who organize content in the way they see fit, both in terms of physical organization (file hierarchies) and logical structures (e.g. taxonomies). The only coordination with remote hosts is through the introduction of links.

  With the present web, the structures to organize the information are not accessible to machines. While links are directly processable (and indeed used by search engines), they carry no meaning to machines other than 'this page is related in some way to that'. This means that well-known search engines such as Google can make no use of this information and at best [26] search the documents for keywords instead. Directories such as Yahoo! circumvent the problem by imposing their own monolithic structure and classify web pages according to that. Besides scalability issues, such effort is contradictory to the original idea of local control over both resources and their organization.

Distributed approaches to Knowledge Management are also advocated by those contend that knowledge, by its very nature, is subjective to a person and local to a community. The emerging discipline of Distributed Knowledge Management (DKM) thus concerns the management of the local processes of knowledge creation and the bilateral exchange of knowledge across autonomous groups [Bonifacio *et al.*, 2002]. Theorists of DKM such as Bonifacio derive DKM from the characteristics of knowledge:

---

[24] For an overview of virtual community issues we recommend the exhaustive collection compiled by the Center for Advanced Learning Technologies (CALT) at INSEAD, available at http://www.insead.fr/CALT/Encyclopedia/ComputerSciences/Groupware/VirtualCommunities/.

[25] Survey of January, 2003. Source: Internet Software Consortium (http://www.isc.org/

[26] It is a different question that search engines often *choose* to ignore available metadata to exclude deliberate attempts to manipulate their ratings.

- Knowledge is subjective. Interpretation of knowledge is intrinsically subjective as it's dependent on the context of the interpreter and his mental schemata. Knowledge as an absolute concept is replaced by the notion of limited, partial interpretations of the world.
- Collective knowledge is a social artifact. Similarly to the previous point collective knowledge only exists, because people share part of their schemas and therefore share some of the understanding. Knowledge is built by a continuous process of negotiating interpretation schemata among people with similar or different perspectives. (With the intent of improving performance and sustaining innovation, respectively.)

These considerations support a strong argument against centralized approaches to KM: as there is no knowledge independent of interpretation and sociality, knowledge can not be simply transferred across communities as any other resource. Instead, knowledge has to be independently managed where it's created and used. The most we can do in terms of management is in coordinating the negotiations between autonomous knowledge sources.

Given this pretext, the most forceful argument for decentralized ontology-based technology is an architectural one. Technological architectures can shape organizational forms and vice versa: organizational forms have great impact on the appropriation of technology [Bonifacio *et al.*, 2002]. The Semantic Web is largely considered as an enabling technology for DKM, because it naturally accommodates a system of autonomous groups, where ontologies are used for representing (part of) a community's interpretation schema. Autonomous groups may evolve their knowledge independently in space and time.

Autonomy needs to be balanced with coordination in order to make interactions possible in cases where groups depend on each other to achieve their goals. Ad-hoc coordination between groups is represented by the merging or mapping of ontologies, a task that can be at least partially automated with today's technology. Knowledge networks are expected to gradually emerge from this bottom-up process of making the connections between the ontologies of single nodes. Note that this network is not static, but itself reconfigures as the underlying ontologies evolve. Since it's invoked on the basis of need for interaction, the structure of this network will reflect the goals, interests etc. of the various groups in the network.

It is worthwhile to note this knowledge network will not be completely flat, but develop a power structure over time. In cases where there is a long term need for interaction within an industry or community, standard ontologies will emerge around the consensus about the domain. Just as there are a handful of websites on the current web embedded connected through a disproportionately large number of links, these ontologies will form the hubs of the Semantic Web and play a distinguished role in mediating between and across domains.

Machine-based interaction without central control promises agility and scalability beyond traditional limits. Here, scalability is defined not so much as the ability of applications to grow in size or geographical extent, but rather to easily extend to a large number of independent organizations or autonomous groups. For the individual groups this brings advantageous economies of scale. As on the present web, the marginal cost of adding communities to the existing infrastructure is minimal. Therefore they can expect faster, if not immediate returns on their investment in joining a knowledge network, as opposed to conventional KM systems that require a large up-front investment [Fensel *et al.*, 2003].

In the following, we discuss the benefits and challenges to the application scenarios introduced before.

6.2.1 Communication in distributed settings

In cases where agents and communities with diverse views of the world would like to share knowledge, they must be able to learn how to understand each other's conceptualizations [Williams, 1999]. Understanding, however, requires communication to be present on the first place. As [Fensel *et al.*, 2003] suggests, the way to break out of the vicious circle of no communication without understanding and no understanding without communication is to accommodate the *process* of repeated, frequent exchanges of meaning and knowledge in the context of particular objects.

The situation of diverse communities connected through the Semantic Web may be paralleled to Distributed Communities of Practice (DCoPs) [Hildreth *et al.*, 2000]. DCoPs are groups of professionals (for example in multinational organizations) who carry out their work in coordination, while separated geographically from each other. In their study of the Museum of Vertebrate Zoology, Star and Griesemer have found two key factors to the success of coordination work in such settings [Star and Griesemer, 1989].

First, there is a need for standardized methods to discipline information obtained by different actors. This means standard methods for the representation of information (or at least a degree of compatibility), the assurance of quality and a common ground in basic tasks related to this information. Second, there is need for what they call *boundary objects*. Boundary objects are artifacts shared by different social worlds, i.e. they cross the boundaries of communities. Yet, they are flexible enough to adapt to different viewpoints and satisfy informational requirements of each different social world. Altogether, boundary objects need to maintain coherence across intersecting social worlds, but allow different interpretations.

Standardized methods for the representation and manipulation of information will be provided by the Semantic Web infrastructure. For communication to work, the Semantic Web will also need to serve as a social infrastructure that provides for the meeting of communities with similar tasks, backgrounds or interests. When people share knowledge, they are not just sharing information: they are also sharing cultural and social references. Similarly, when people seek knowledge, they are not just seeking information, but information grounded in, and carrying different meanings to different social communities [Mantovani, 1996].

The second factor, boundary objects are played by the ontologies and metadata provided by communities connected through the Semantic Web. With information derived from references to the shared context, e.g. elements of the task or external references, the ontologies may be mapped partly, which allows partners to obtain a partial understanding of the other's view on the context. Within certain limits, this will allow the community to take another perspective of the world and even adapt his local theory accordingly, if it's deemed beneficial to do so. Altogether, multiple languages and partial interpretations will replace a single ontology and full commitment; as there is no ontology without a community, we cannot expects agents from diverse social worlds to fully share the commitment behind a conceptualization. Fortunately, that is not required: shared understandings will need to be developed only to the point that is required by the particular task at hand.

In summary, the Semantic Web will provide the space where the dynamic, codependent processes of communication and understanding will take place. Ontologies and metadata formulated with those ontologies will serve as the boundary objects; partial understandings will be developed over them through the mapping of ontologies based on a shared context. These exchanges will mimic the rumblings of a group of people with diverse cultural backgrounds trying to evolve their understanding towards greater efficiency in problem solving.

The application of these methods will also bring a number of new challenges. Much like on today's WWW, we expect no authority on the Semantic Web above the level of autonomous groups, which represent the nodes of the network. To a lesser extent, these concerns are relevant to corporate applications of Semantic Web technology as well:

- **Motivation.** While hardly an issue in a controlled environment, the question of motivation will play an important role in the usage of Semantic Web applications. Starting from inducing a network effect (getting enough nodes to join the network in order to make the network attractive enough for others to join), to the question of maintenance of information, motivation will need to take the place of authoritative control.
- **Trust.** As Karl-Erik Sveiby, one of the founding fathers of Knowledge Management phrased it, "Trust is the bandwidth of communication." As on the present web, the amount of trust in the source of knowledge will put an upper limit on its usefulness. Disruptive knowledge or 'semantic spam' will also need to be dealt with.
- **Privacy.** Privacy and the protection of intellectual assets is already a major concern for industries that have seen much of these assets bartered away in P2P content networks such as Napster, Kazaa etc.
- **Uncertainty.** Unless there are enforced guarantees, nodes of the knowledge network might come and go at the pace at which today's websites appear, relocate or disappear.

6.2.2 Integration in distributed settings

Albeit linking takes place at a different level, as in the standalone case the integrative force of ontologies is a direct outcome of their linking power. While ontologies of autonomous groups are used to integrate resources within their closed environments, a global ontology or the many ontologies that result from mappings integrate resources across autonomies into a single knowledge network. This network connects people to other people, resources or services, wherever they reside. Again, tapping into this network is much easier as ontologies help to formulate queries, aid browsing, support analysis of resources etc.

The benefits and challenges are similar to the ones mentioned in the previous section. As there is no authority above the level of single nodes, nodes may join or leave the network or reconfigure themselves as they see fit. This dynamics can be especially useful where a changing environment or the nature of the task demands such loose coupling and constant rearrangement. This is the case for virtual communities or organizations that form around a specific task and for a limited duration. Dynamic knowledge networks are also expected to fare better in case of reorganizations within a company, since they are easily reassembled [Bonifacio *et al.*, 2002; Williams, 1999].

6.2.3 Reasoning in distributed settings

Again, this is the most complex application scenario in terms of the supporting ontologies and applications, although this scenario is by far the most intriguing one as well.

The potential attributed to reasoning in a distributed setting rests on the single fact that aggregated knowledge is potentially far more powerful than the sum of its parts. By tapping into a distributed system of knowledge bases, we are potentially able to draw conclusions that one could not base on any of the single bodies of knowledge. This is the real vision of the Semantic Web as exposed by Tim Berners-Lee's popular article in the Scientific American [Berners-Lee *et al.*, 2001]: a web where agents (distributed application components) that live on the knowledge network described before work together and combine knowledge from various sources to carry out such complex tasks as scheduling appointments for us, doing our Christmas shopping, etc.

While processing knowledge, agents will need to deal with operate in an open environment where the concepts of absolute truth, total knowledge and total provability are meaningless [Berners-Lee *et al.*, 1999]. What remains is at best limited knowledge that is only true with a certain probability based on the trust we put in the source we used to derive it. Semantic applications will also need to work robustly in face of inconsistencies.

The quality of ontologies will also degrade through the extensive use of automated methods in ontology acquisition, translation and mapping. While this seems to be a technical issue at first sight, we will argue that the quality of these methods will never reach the level that can be attained by manual approaches employed in small scale applications and narrow domains. Degradation will be especially noticeable for knowledge with a large sharing scope, such as the common sense required to solve the tasks in Berners-Lee's scenario. We will return to these issues in Section 7.

7 Discussion

The goal that we have set forth in this paper was to provide a synthesis of ontologies and their most significant application area, the business discipline of Knowledge Management. We have been motivated in this effort as we see the ontology and KM communities developing in parallel, leaving potential synergies unexploited despite a match between needs and potential solutions. Without bringing these communities closer we feel that the ontology field would be in clear danger of producing tools and techniques that prove to be a mismatch to real needs, while the KM community would miss out on the opportunities presented by the appropriate use of ontologies.

In the previous sections of this paper we have built the Semantic Web matrix, a two-dimensional classification of ontology-based Knowledge Management applications. The first dimension distinguished three layers of ontology functionality: a lightweight use of ontologies for communication, integration scenarios using metadata and complex reasoning systems that rely on highly formal knowledge. The second, centralization and control dimension of the matrix ranged from stand-alone applications under the control of a single authority to systems with distributed application components (MAS, Grid, Semantic Web services).

In the following, we discuss what the message of this matrix is with regard to ontologies for Knowledge Management and KM for ontologies.

7.1 Ontologies for Knowledge Management

Through our survey of the KM market in Section 3, we have seen that the KM market is indeed ready for the adoption of ontologies, even if this would happen without carrying over the name itself. Nevertheless, we are hopeful that by educating the KM community ontologies can establish a name of their own as conceptual models that go beyond simple taxonomies. KM systems would benefit from the scenarios enabled by more complex forms of knowledge and could profit from the ongoing standardization of the technology (e.g. the OWL web ontology language).

Unfortunately, at the moment there is still a vicious circle, where ontologies are relatively unknown due to the lack of convincing applications and applications are missing due to the lack of industry buy-in. To break this circle, ontologies need to be explained and made convincing enough for the Knowledge Management community. In this paper we provided advice on how to argue for ontology-based applications in terms familiar to Knowledge Management professionals and in a way that fits the current trends in KM technology. This step-by-step approach also means that instead of the grand vision of the Semantic Web, such argument should focus on the organizational or inter-organizational use of the technology. As opposed to the global Semantic Web which suffers from a bootstrapping problem, enterprize applications of ontologies hold the more immediate benefit due to the smaller scope and the easier acquisition and control of metadata.

We would also add the lesson that technology supplying the KM market should be aware of the complex issues surrounding Knowledge Management. For a technology-driven community, it is particularly important to realize the delicate interactions between technology, people and organizations, the setting in which the technology is applied. This context puts serious limitations on the reach of technology, which makes it all the more important for the ontology community to create solutions that are at least aware of these issues. For example, applications should be made flexible enough to be adapted to the needs of different users and task environments. Explicit user and process descriptions managed by the middleware should be used to tailor the application environment to the interest or task of a specific user, his strategy for problem solving etc. Methodologies should also put more emphasis on the soft issues of introducing a knowledge system, for example the possible gratifications that could motivate people to provide and maintain community knowledge voluntarily.

The already existing showcases should be used more effectively in promoting ontology technology at the regular forums and events of the KM community. Ideally, this would lead to a two-way discussion, where the industry feedback and experience could help us to arrive at a more focused research agenda. This issue should be taken up more actively by research policy makers and network organizations both in the EU and outside.

7.2 Knowledge Management for Ontologies

One of the outcomes of the study of Knowledge Management and the construction of the Semantic Web Matrix is an increased awareness that ontology research will need to pay more concern in the future to the *social* nature of knowledge.

Ontologies, as other forms of knowledge, are socially constructed objects. This means that they are limited in their reach to settings outside the community in which they were conceived. How large this sharing scope is depends on the specificity and dynamics of the knowledge concerned as well as the trust attributed to it by other communities.

The particular contradiction between the formality, stability, and sharing scope of knowledge has been studied by Elst and Abecker [van Elst and Abecker, 2002]. While the authors discuss agents in information systems, their argument can be just as well understood in real terms. For example, one only has to think of how highly technical books are limited in their reach to the experts familiar with the formalisms and the details of the subject matter (formality[27] vs. sharing scope). Also, books that describe a narrow area of expertise in much detail are likely to have frequent revisions (formality vs. stability). Frequent revisions to the subject matter of a book also limit the scope of the knowledge it contains, as some people will have difficulty to adapt their own knowledge and follow up with the field (stability vs. sharing scope).

The contradiction of formality and sharing in ontology-based Knowledge Management is vividly demonstrated by our Semantic Web Matrix: applications towards the upper right corner are increasingly difficult to implement in practice due to the difficulty of moving along both axis simultaneously (see Figure 8.) Moving along the Semantic Dimension requires increasing formality, while advancing the Web dimension leads to more dynamics (loss of control, uncertainty) and larger sharing scopes. In the light of the above theory these are contradictory properties of knowledge.

We note that these contradictions can be readily observed in the ontology domain. While heavyweight ontologies are powering focused, narrowly scoped ‘expert’ knowledge systems, lightweight applications of RDF (in particular FOAF and RSS) are behind the overwhelming majority of metadata on the public web.[28] The tension between scope and formality is also reflected in the difficulty of moving ahead in the search for better means of automated ontology learning and mapping. Much like humans, highly trained methods that work well in one domain produce unacceptable results when moved to another, unknown domain.

---

[27] While Elst and Abecker mean training in the formalism itself, a shared social context is also required for understanding. While the formality of ontologies constrains the interpretation of their symbols, the creation of meaning (i.e. the interpretation of symbols in context) remains the task of a human agent, whose knowledge cannot be detached from his personal and social existence.

[28] News and personal information have made the current web and this pattern is likely to be repeated on the Semantic Web. The Friend-Of-A-Friend (FOAF) ontology allows to create a personal profile. FOAF also provides a mechanism to link one’s profile to the profiles of his friends, creating a lightweight social network in RDF format. The RDF Site Summary (RSS 1.0) ontology is used to describe news items of any kind, which allows applications for aggregating and syndicating news from various sources.

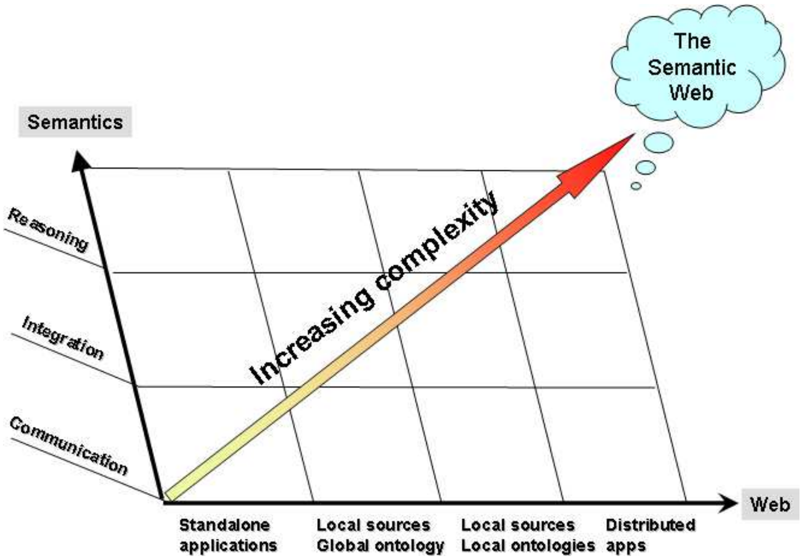


Figure 8: Complexity increases rapidly as we try to move along both axis at the same time due to the difficulty of maintaining formality in decentralized settings.

The ramifications for the Semantic Web require further investigation. Here we would only postulate that contingency will be provided by the natural clustering of the Semantic Web in overlapping domains of knowledge communities. As illustrated in Figure 9, due to the above mentioned constraints a Semantic Web of agents is likely to replicate the structure of human societies where distant communities are woven together by a selected number of hubs, with the special capability of moving at ease between different communities, translating and carrying messages across. However, for reasons of efficiency in problem-solving, the majority of the agents will have to be constrained to building strong ties with a limited set of peers in their immediate neighborhood.

What the Semantic Web will then need is an infrastructure that accommodates the social constructs behind knowledge: an explicit modelling of peers and communities, linked by a rich set of social relations. Making our information systems aware of this context will support the human process of interpreting information: sociality is a precondition for something to have meaning [Bonifacio *et al.*, 2002] as the ability to interpret an expression hinges on grounding symbols to the same objects by encoder and interpreter. Similarly, the relevancy of knowledge is determined by a shared task or interest of those involved. Lastly, without the social notions of trust and legitimacy, agents on the Semantic Web will not be able to make intelligent decisions about which knowledge to allow to enter their reasoning processes. As a result, 'semantic spam' in the form of misleading, inconsistent or incorrect knowledge will overburden Semantic Web agents as spam now litters our mailboxes every day.

We hope that through our analysis ongoing and future research projects that promote ontologies in the name of Knowledge Management will become more aware of the peculiar state of the art. An interactive approach to ontology-based KM will once

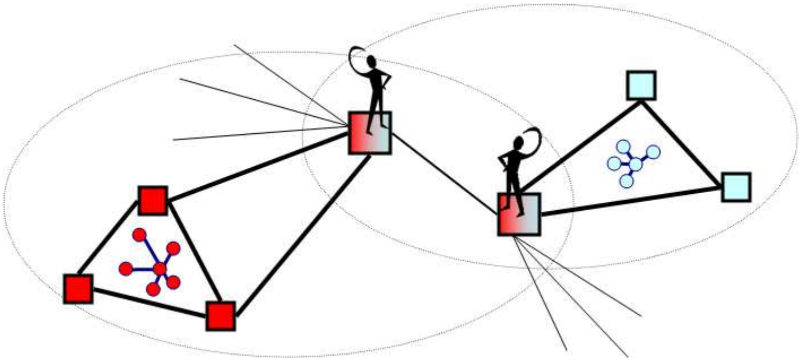


Figure 9: The architecture of the Semantic Web should reflect the social aspects of communication and learning.

lead hopefully to a general acceptance of ontologies as another tool in the toolkit of knowledge engineers and managers. As for the Semantic Web, there will be no final announcement or fanfare to let us know when we reach it. And in many senses, we are there already . . . .

Acknowledgments. The authors acknowledge the support for this work by the European Commission, particularly in the context of the EU-IST Project SWAP [swa,], the thematic network OntoWeb [OntoWeb,], and the European Network of Excellence KnowledgeWeb [KnowledgeWeb,]. Many individuals from these and other projects contributed to this work, via direct discussions, or willingness to participate in surveys or interviews. We are especially grateful to Frank van Harmelen (Free University, Amsterdam) and Steffen Staab (Institute AIFB, University of Karlsruhe) for their comments to the manuscript.

References

[Abecker and van Elst, 2003] Andreas Abecker and Ludger van Elst. *Handbook on Ontologies*, chapter 22: Ontologies for Knowledge Management, pages 436–454. International Handbooks on Information Systems. Springer Verlag, 2003.

[Ackoff, 1999] Russell L. Ackoff. *Ackoff's Best: His Classic Writings on Management*. John Wiley & Sons, 1999.

[Berners-Lee *et al.*, 1999] Tim Berners-Lee, Mark Fischetti, and Michael L. Dertouzos. *Weaving the Web : The Original Design and Ultimate Destiny of the World Wide Web by its Inventor*. Harper San Francisco, 1999.

[Berners-Lee *et al.*, 2001] Tim Berners-Lee, James Hendler, and Ora Lassila. The Semantic Web. *Scientific American*, May 2001.

[Bonifacio *et al.*, 2002] Matteo Bonifacio, Paolo Bouquet, and Paolo Traverso. Enabling Distributed Knowledge Management: Managerial and Technological Implications. *Novatica and Informatik/Informatique*, III(1), 2002.

[Borst *et al.*, 1997] W. N. Borst, J. M. Akkermans, and J. L. Top. Engineering ontologies. *International Journal of Human-Computer Studies*, 46:365–406, 1997.

[Buzan and Buzan, 1996] Tony Buzan and Barry Buzan. *The Mind Map Book: How to Use Radiant Thinking to Maximize Your Brain's Untapped Potential*. Plume, 1996.

[Cost *et al.*, 2002] Scott Cost, Tim Finin, Anupam Joshi, Yun Peng, Charles Nicholas, Ian Soboroff, Harry Chen, Lalana Kagal, Filip Perich, Youyong Zou, and Sovrin Tolia. ITtalks: A Case Study in the Semantic Web and DAML+OIL. *IEEE Intelligent Systems*, 17(1):40–47, January, February 2002.

[Davenport and Prusak, 1998] Th.H. Davenport and L. Prusak. *Working Knowledge*. Harvard Business School Press, Boston, MA, 1998.

[Dean *et al.*, 2004] Mike Dean, Dan Connolly, Frank van Harmelen, James Hendler, Ian Horrocks, Deborah L. McGuinness, Peter F. Patel-Schneider, and Lynn Andrea Stein. OWL Web Ontology Language Semantics and Abstract Syntax. Technical report, World Wide Web Consortium (W3C), February 2004.

[Dieng *et al.*, 1999] Rose Dieng, Olivier Corby, Alain Giboin, and Myriam Ribiere. Methods and Tools for Corporate Knowledge Management. *International Journal of Human-Computer Studies (IJHCS)*, 51:567–598, 1999.

[Drucker, 1993] P.F. Drucker. *Post-Capitalist Society*. Butterworth-Heinemann, Oxford, UK, 1993.

[Dublin Core, 1999] Dublin Core Metadata Element Set, Version 1.1: Reference Description. http://dublincore.org/documents/dces/, July 1999.

[Ewalt, 2002] David Ewalt. The Next Web. *InformationWeek*, October 2002.

[Fensel *et al.*, 2003] Dieter Fensel, Steffen Staab, Rudi Struder, Frank van Harmelen, and John Davies. *Towards the Semantic Web: Ontology-Driven Knowledge Management*, chapter 14: A Future Perspective: Peer-2-Peer Enabled Semantic Web for Knowledge Management. John Wiley & Sons, 2003.

[Godbout, 1999] Alain J. Godbout. Filtering Knowledge: Changing Information into Knowledge Assets. *Journal of Systemic Knowledge Management*, January 1999.

[Gordijn and Akkermans, 2001] J. Gordijn and J.M. Akkermans. Designing and evaluating e-Business models. *IEEE Intelligent Systems*, 16(4):11–17, 2001.

[Gordijn and Akkermans, 2003] J. Gordijn and J.M. Akkermans. Value-based requirements engineering: Exploring innovative e-Commerce ideas. *Requirements Engineering*, 8:114–134, 2003.

[Group, 2002] Delphi Group. Taxonomy and Content Classification. A Market Milestone Report. Whitepaper, 2002.

[Gruber, 1993] Tom Gruber. A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, 5(2):199–220, 1993.

[Hildreth *et al.*, 2000] Paul Hildreth, Chris Kimble, and Peter Wright. Communities of practice in the distributed international environment. *Journal of Knowledge Management*, 4(1):27–38, 2000.

[Iosif and Mika, 2002] Victor Iosif and Peter Mika. EnerSearch Virtual Organization Case Study: Evaluation Document. On-To-Knowledge Deliverable 29, 2002.

[Iosif *et al.*, 2003] Victor Iosif, Peter Mika, Rikard Larsson, and Hans Akkermans. *Towards the Semantic Web: Ontology-Driven Knowledge Management*, chapter 13: Field Experimenting With Semantic Web Tools In A Virtual Organization. John Wiley & Sons, 2003.

[KnowledgeWeb,] The EU KnowledgeWeb Network of Excellence (NoE) project. http://knowledgeweb.semanticweb.org.

[Lassila and Swick, 1999] Ora Lassila and Ralph R. Swick. Resource Description Framework (RDF) Model and Syntax Specification. Technical report, World Wide Web Consortium (W3C), February 1999.

[Leger, 2002] Alain Leger. Successful Scenarios for Ontology-based Applications V1.0. OntoWeb Deliverable 2.2, 2002.

[Lueg, 2002] Christopher Lueg. Knowledge management and information technology: Relationship and perspectives (introduction to the special issue). *Novatica and Informatik/Informatique*, III(1), February 2002.

[Mantovani, 1996] Giuseppe Mantovani. Social context in HCI: A new framework for mental models, cooperation, and communication. *Cognitive Science*, 20(2):237–269, 1996.

[Manville and Foote, 1996] Brook Manville and Nathaniel Foote. Harvest Your Workers' Knowledge. *Datamation*, July 1996.

[Motta *et al.*, 2003] Enrico Motta, John Domingue, Liliana Cabral, and Mauro Gaspari. IRS-II: A Framework and Infrastructure for Semantic Web Services. In John Mylopoulos Dieter Fensel, Katia P. Sycara, editor, *Proceedings of the 2nd International Semantic Web Conference (ISWC2003)*, volume 2870 of *Lecture Notes in Computer Science*. Springer Verlag, 2003.

[Newell, 1982] Allen Newell. The knowledge level. *Artificial Intelligence*, 18(1):87–127, January 1982.

[Nonaka and Takeuchi, 1995] Ikujiro Nonaka and Hirotaka Takeuchi. *The Knowledge-Creating Company*. Oxford University Press, 1995.

[OntoWeb,] The EU OntoWeb Thematic Network (TN) project. http://www.ontoweb.org.

[Peter Mika and Daniel Oberle and Aldo Gangemi and Marta Sabou, 2004] Peter Mika and Daniel Oberle and Aldo Gangemi and Marta Sabou. Foundations for Service Ontologies: Aligning OWL-S to DOLCE. In *Proceedings of the 13th International World Wide Web Conference (WWW2004)*. ACM Press, 2004.

[Reimer *et al.*, 2003] Ulrich Reimer, Peter Brockhausen, Thorsten Lau, and Jacqueline R. Reich. *Towards the Semantic Web: Ontology-Driven Knowledge Management*, chapter 12. Ontology-Based Knowledge Management at Work: The Swiss Life Case Studies. John Wiley & Sons, 2003.

[Schreiber *et al.*, 1999] Guus Schreiber, Hans Akkermans, Anjo Anjewierden, Robert de Hoog, Nigel Shadbolt, Walter van de Velde, and Bob Wielinga. *Knowledge engineering and management. The CommonKADS Methodology.* MIT Press, 1999.

[Schreiber *et al.*, 2002] A. Th. Schreiber, I. Blok, D. Carlier, W. van Gent, J. Hokstam, and U. Roos. A mini-experiment in semantic annotation. In Ian Horrocks, editor, *International Semantic Web Conference 2002*, Lecture Notes in Computer Science. Springer Verlag, 2002.

[sem,] Whitepaper: National Security and Intelligence. http://www.semagix.com.

[Smith and Welty, 2001] Barry Smith and Christopher Welty. Ontology: Towards a New Synthesis. In *Proceedings of the International Conference on Formal Ontology in Information Systems (FOIS2001)*. ACM Press, 2001.

[Star and Griesemer, 1989] Susan Leigh Star and James Griesemer. Institutional Ecology, Translations, and Coherence: Amateurs and Professionals in Berkeleys Museum of Vertebrate Zoology, 1907-1939. *Social Studies of Science*, (19):387–420, 1989.

[swa,] The SWAP project (EU-IST-2001-34103). http://swap.semanticweb.org.

[Uschold and Jasper, 1999] Mike Uschold and Robert Jasper. A Framework for Understanding and Classifying Ontology Applications. In *Proceedings of the IJCAI99 Workshop on Ontologies and Problem-Solving Methods(KRR5), Stockholm, Sweden*, 1999.

[van Elst and Abecker, 2002] Ludger van Elst and Andreas Abecker. Ontologies for information management: balancing formality, stability, and sharing scope. *Expert Systems with Applications*, 23(4):357–366, November 2002.

[van Engers, 2001] Tom Maarten van Engers. *Knowledge Management: The Role of Mental Models in Business Systems Design*. PhD thesis, Vrije Universiteit, Amsterdam, 2001.

[Williams, 1999] Andrew B. Williams. The Role of Multiagent Learning in Ontology-based Knowledge Management. In *American Association for Artificial Intelligence, AAAI Spring Symposium, Stanford University*. AAAI Press, 1999.

[Wilson, 2002] Tom Wilson. The nonsense of ’knowledge management’. *Information Research*, 8(1), October 2002.