D21 Successful Scenarios for Ontology-based Applications v1.0

OntoWeb
Ontology-based information exchange for knowledge management and electronic commerce
IST-2000-29243

Date: 15th January, 2002

Identifier	**Deliverable 2.1**
Class	**Deliverable**
Version	**1**
Version date	**15-01-2002**
Status	**Draft**
Distribution	**Public**
Responsible Partner	**France Telecom R&D**

Project ref. no.	*IST-2000-29243*
Project acronym	OntoWeb
Project full title	Ontology-based information exchange for knowledge management and electronic commerce

Security (distribution level)	Public deliverable
Contractual date of delivery	31 November 2001
Actual date of delivery	January 16, 2002
Deliverable number	D2.1
Deliverable name	Successful scenarios for ontology-based applications
Type	Report
Status & version	V1.0 Final
Number of pages	59 pages
WP contributing to the deliverable	WP2
WP / Task responsible	WP2
Main contributors	Hans Akkermans, Michael Brown, Jean-Marc Bouladoux, Rose Dieng, Ying Ding, Asunción Gómez-Pérez, Siegfried Handschuh, Anthony Hegarty, Andreas Persidis, Rudi Studer, York Sure, Brigitte Trousse, Valentiina Tamma
Editor(s)	Alain Léger
EC Project Officer	Hans-George Stock
Keywords	Ontology-based applications, Successful Business Scenarios, Best Practices, Guidelines for implementers, Key findings
Abstract (for dissemination)	This deliverable presents a first version of a series of 5 documents, to be delivered each 6 months and whose target is to give guidelines for the application of Knowledge IT to practitioners in the field of - in large - E-Commerce and Knowledge Management. This series of documents is primarily aimed at anyone who is involved in the process of designing, building and managing Knowledge-Based systems for the field mentioned above. It should assist anyone in the industry or the commercial sectors to evaluate how these emerging technologies are suitable to answer practical needs. Furthermore, each successful retained scenario is illustrated by practical examples.

Document History

Version	Date	Comments and Actions	Status
V0.1	**24/11/2001**	**First draft version edited by co-ordinator**	**Draft proposal**
V0.2	**02/12/2001**	**Second draft version edited by co-ordinator**	**Draft proposal**
V0.3	**06/12/2001**	**Third draft version co-edited with contributions**	**Draft proposal**
V0.4	**04/01/2002**	**Revised version for completion by all actors**	**Draft proposal**
V0.5	**10/01/2002**	**Revised version for completion by all actors**	**Draft proposal**
V0.6	**10/01/2002**	**Integration of W2 members contributions**	**Draft proposal**
V0.7	**14/01/2002**	**Integration of contributions**	**Draft proposal**
V0.8	**15/012002**	**Integration of contributions**	**Draft proposal**
V0.9	**16/01/2002**	**Finalizations**	**Draft proposal**
V1.0	17/012002	**Final proofing**	Final

Acknowledgements

The Editor of D21 would like to thank specifically the following main contributors:

- **York Sure, Rudi Studer** and **Siegfried Handschuh** (University of Karlsruhe, Germany) for the Knowledge Management part and the Semantic portal part ;
- **Andreas Persidis** (Biovista) for the Information Retrieval part and especially for the continuous support of this work ;
- **Rose Dieng** and her team (INRIA Sophia Antipolis, France) for the Knowledge Management part ;
- **Martin Bryan** (The SGML Center, UK) for the E-Commerce part ;
- **Mike Brown** and **Asuncion Gomez-Perez** for the global effort to edit D12 (“Business Scenarios”) in harmony with D21 (that deliverable : “Successful Scenarios and Guidelines”) ;
- **Ying Ding** for the constant effort to provide many materials for the completion of the D21 ;
- **Hans Akkermans** for the information provided for the E-Commerce section and the global effort to help in providing right resources within the SIG4 (Applications)
- **Jean Marc Bouladoux** (Odos Strategy sa Paris) for contributions on Evaluation ;
- Important Contributions on KBS Evaluation (so rare) based on **Tim Menzies** papers and references
- Contribution from Best Practices LLC (**Susan Silverstein**, Sales Executive)
- **Many others contributors** (always referred in the References section) to the **E-work and E-Commerce 2001** event held in Venice (15-17 October 2001) and sponsored by the European Initiative in **New Methods of Work and Electronic Commerce**, also known as Key Action II in the European Information Society Technologies (IST) programme.

Executive summary

This deliverable is the **first of a series of 5 documents** to be delivered on a **6 monthly basis** and whose target is to report **Best Practices and to give Guidelines for the application of Knowledge IT** to practitioners in the field of **-** in large **-** E-Commerce and Knowledge Management.

This series of documents **is primarily aimed at anyone who is involved in the process of designing, building and managing Knowledge-Based systems** for the field mentioned above. It **should assist anyone in the industry or the commercial sectors to evaluate how these emerging technologies are suitable to answer practical needs**. Furthermore, each successful retained scenario is illustrated by practical examples.

Its **global scope** includes

- Presentation of application problems
- Application Functional needs and Main Business Use Cases
- Evaluation of Ontology-Based Applications
- Successful Business scenarios and Guidelines
- Tools for Ontology-based applications

This first deliverable of the series, namely **D21, is aiming at**:

- Examination of the state of the art on ontology-based applications and categorising of the most representative, mature and successful ontology-based application clusters
- Identification of killer applications in each cluster
- Guidelines to assist an organization to classify applications so as to follow an appropriate implementation methodology

Key Findings, Key Challenges and Prioritised Lessons Learned at this stage (D21)

At this stage of the Analysis we would like to clearly make a distinction between two kinds of evaluation:

1. Internal properties level evaluation of Knowledge Based system (e.g. syntactic anomalies : tautologies, cycles, …) as usually performed by the verification community
2. Service and Business level evaluation (e.g. utility, return-on-investment, economic suitability, …)

We claim that D21 must only address the level of Service and Business (level 2).

Next Step **D22 (15 May 2002)**
The second deliverable of the series, namely D22, will be aiming at:

- Identification of the most suitable techniques that may be applied to each cluster of ontology-based applications
- Provision of Guidelines to assist an organisation on **what** techniques could be applied for a given cluster of application

Table of contents

1 SCOPE

1.1 Introduction

Knowledge-based technology is increasingly being applied to a very large set of applications. To name the most prominent ones we can enumerate: *Knowledge Management*, *E-Commerce*, *Natural Language processing* and *Machine Translation*, *Information Retrieval*, *Integration of Heterogeneous Information*, *Intranet Community Portals*, etc. **Interest in the evaluation of the Knowledge-based technology** is growing due to the **increased maturity of the field** and due to **business pressure** for the evaluation of the **measurable costs and benefits of that technology**. Being **ahead of the competition** involves employing **key technologies** and **capturing essential guidelines and services**. Unfortunately, as a matter of facts, quantitative studies on that field are **relatively rare**.

The major goal of the OntoWeb WP2 group is to issue **Best Practices and Guidelines** in providing **benchmarking methodologies based on the performance characteristics** of various Ontology Based systems designed for different **classes of applications**.

Each Best Practice and Guideline document will describe

- the class of system or services being addressed;
- discuss the performance characteristics that are pertinent to that class;
- clearly identify a set of metrics that aid in the description of those characteristics;
- specify the methodologies required to collect said metrics;
- and lastly, present the requirements for the common unambiguous reporting of Best Practice results.

Benchmarking is the process of seeking out**, studying and quantifying the best practices that produce superior performance**. The traditional metrics-focused approach must be supplement with an analysis of **why and how** practices produce better results. It should help **understanding strengths and weaknesses** and can be supplemented by a road map for improvement (companion document D11). The determination of benchmarks allows one to make a direct comparison of off-the-shelves bundle solutions or its own solution against competitors. Any identified gaps in the comparative study are potential improvements
Another form of benchmarking includes "process benchmarking" generally higher-level and less numbers-intensive than metrics. These studies demonstrate how top-of-classs application accomplish the specific process in question.
Best practices are documented strategies and process employed by top-of-class Applications. One given application can not claimed to be top-of-class in every area **-** such application does not exist **-** so in D21 and series, we are capturing the best practices of the top best applications in a given class.
We collected this information from a variety of sources available from the OntoWeb Network as a whole. This can take the form of interviews, experiences, surveys, publications, books, magazines, libraries and Internet.
This collection of information is monitored and reported in a companion document series D12[1] (Technological Road Map of Business Scenarios), D12 serve as **a means of assimilating information**

[1] The relationship between deliverables D1.2 and D2.1 is as follows:

- **Common Web Portal –** Is used to gather (anecdotal) information from the community useful for both deliverables
- **Deliverable D1.2 (Business Scenarios)** distils from this information and other sources the reasons *Why* and under *What Conditions* a commercially viable ontology application can be deployed. Deliverable D1.2 will also offer some predictions as to future commercial trends.
- **Deliverable D2.1 (Best Practices and Guidelines)** goes further in attempting to provide a concrete set of guidelines as to what needs to be achieved in order to deploy ontology solutions, i.e. this deliverable describes the *How to* construct as successful ontology based application.

about the diversity of commercial ontology projects currently in development or already deployed. A **table of generic parameters characterizing key business scenarios / applications** is provided.

This first deliverable of the series, namely **D21, is at least aiming at**:

- Examination of the state of the art on ontology-based applications and categorizing of the most representative, mature and successful ontology-based application clusters
- Identification of killer applications in each cluster
- Guidelines to assist an organization on **which** cluster a given application could be classified

1.2 Business Models for the Semantic Web?[2]

Here is given the kind of related questions that the study currently pursued in the D21 series is trying to answer.

- However, business-wise, the Semantic Web is only at its very beginnings. Many aspects are still unclear:
    - What are the specific problems for which identifiable, reachable purchasers will buy solutions that are based on Semantic Web technology, and which cannot reasonably be implemented any other way? As an example, consider eBay for today's web: eBay clearly fills a need, and could not exist in any other medium. In other words, what makes a solution based on the Semantic Web "10 times" better than what we have today?
    - What will be the emerging "business ecosystem" around the Semantic Web? What are reasonable product/solution categories? Some potential categories are: Ontology modeling tools, ontology servers, servers to execute logical assertions, equivalence/ontology transformation servers, etc. Is the Semantic Web going to be a client-server system, or a peer-to-peer system with symmetric protocols? Is there going to be a stack of technologies on top of each other, or even a knowledge stack?
    - Is the Semantic Web primarily an opportunity for upstarts with disruptive technologies, or an opportunity for established companies to provide more value to their existing customer as part of regular product upgrades?
    - How many of the technologies currently being developed under the banner of the Semantic Web are "real" from a business perspective? How many more do we need to be able to implement Whole Products for mainstream users?
    - Will the compelling reason to buy Semantic Web technologies be primarily one of increased revenue or lowered cost for the customer? Or one that is non-monetary in case of consumer adoption?
    - Advertising as a business model for the Semantic Web is clearly not going to work, as the Semantic Web is all about communication between computers, who don't care about advertisements. What are the dominant business models going to be?
    - Even more fundamentally: are we going to see adoption first in the business markets or in the consumer markets? How are we going to solve the chicken-and-egg-problem between information/knowledge providers (who will not do so unless there are users) and information/knowledge users? (who will not make investments in software, education, change of behavior unless there is a Semantic Web out there already)
    - who are going to be the primary producers of Semantic Web content? The case can be made that many existing information providers on the web have strong incentives against providing their content in machine-processable form, as that would lower customer switching costs to competitors If it is not them, who else is it going to be, and how do they assemble critical mass?
- Is the adoption of the Semantic Web going to occur because of:
    - social dynamics (one can argue that today's web initially was adopted not for business reasons, but because it was "cool" in its initial user segment)
    - the increased availability of XML-based data sources
    - the need for enterprise, and inter-enterprise application integration
    - the supply of knowledge brokers/online brokers for business
    - web services

[2] Excerpt from International Semantic Web Working Symposium, Stanford University, CA, USA August 1, 2001

    - the need for machine to machine communication?
- There are many human factors that need to be considered, such as:
    - Why will people annotate their web content, in particular if they do not benefit directly themselves?
    - How many content authors are even qualified to use technologies such as ontologies? How could one grow that number? Or is our current technology just far too hard to use?
    - What kinds of authoring tools are necessary to make this process simple? Do they have to be domain-specific, and if so, how could they be built and distributed efficiently by software companies?
    - How many people are really qualified to build ontologies? How could we foster ontology reuse?
    - How could the Semantic Web piggyback on existing work processes in order to reduce cost of Semantic Web content creation?

1.3 Terminology and Acronyms

We list here some of the terms and acronyms used throughout this document that refer to ontology based applications:

Ontology application scenario : A scenario is an abstract use case for a class of similar applications. Application means a system or a process that makes use of or benefits from the ontology. It describes a particular situation in which an ontology is put to use for some specific purposes.

Actors : Each scenario involves a set of actors. Each actor represents a role that a person or application may play. A person or application may play more than one role in a scenario.

Application User : the role of the user of the application

Benchmarking : Benchmarking is the process of seeking out, studying and quantifying the best practices that produce superior performance. The traditional metrics-focused approach must be supplement with an analysis of why and how practices produce better results. It should help understanding strengths and weaknesses and can be supplement by a road map for improvement. The determination of benchmarks allows one to make a direct comparison. Any identified gaps are improvement areas.

Metrics: It gives numerical standards against which a client's own processes can be compared. Metric benchmarks are of the form:

- Precision and Recall is better than 90% for Ontology-based matching
- Ontology maintenance Cycle time is less than 2 weeks

These metrics are usually determined via a detailed and carefully analysed survey or interviews. Applications developers are then able to identify shortcomings, prioritise action items, and then conduct follow-on studies to determine methods of improvement.

2 Why evaluate Ontology based applications?

2.1 Motivation

In the increasingly competitive knowledge-intensive economy, the search for competitive advantage and the necessity to optimally allocate scarce resources, creates a pressing need for evaluating the ROI as well as the contribution to value creation of applications. This is especially true of IT and knowledge-intensive applications for which this contribution is not always clear and often difficult to justify.

The semantic-web is a relatively new concept whose stated vision is generally accepted by industry, even though there are as yet few successful high-profile cases that support the theory and demonstrate tangible results. However, in order to avoid mistakes that have been made in the past with other leading edge, high profile technologies such as AI, it is essential that the semantic web and the ontology communities devise evaluation methodologies that they will be able to apply from the initial stages of their effort. There are two main reasons why this is useful:

- Management of expectations: As with all new things it is always wise to "promise less and deliver more". Evaluation of ontology-based systems will indicate the scope and limitations of such applications thus allowing developers to understand the capabilities of the technology and deployment time frames and end users to be educated about system potential.

- Guidance of the development and implementation process: evaluation of applications is useful to developers since it provides information and feedback that can be used in their further development.

How are we to assess all the Knowledge Engineering Technology being reported in the Semantic Web literature or advertised on the web by popping-up very young companies ? How can we make them transferable to the Industrial and Commercial community? And above all, how can we make it beneficial for and transparent to the final user as any today ordinary service. We should carefully assess superlative claims usually made about new technologies and especially in the very burgeoning area of Semantic Web technology.

For example in the KA literature, many of the claims in the PSM literature are not supported by the currently available empirical evidence [Menzies, 1997], [Menzies, 1998].

Clearly, we need some better method than trusting the glowing reports available from the promoters of these Knowledge technology or embedded solutions. Even if these authors are expert in their fields, they are very often unable to perform objective expert evaluations. Cohen compares this situation to a parent gushing over the achievements of their children and comments that...

> *What we need is not opinions or impressions, but relatively objective measures of performance.* [Cohen, 1995].

Taking into consideration the wide spread of the potential use of the Knowledge Engineering Technology in applications or solutions, the task of defining objective criteria and methodology to assess Semantic Web Solutions seems extremely challenging.

2.2 Basis for a Generic Evaluation Framework of Ontology based Applications[3]

In this first deliverable (D21) we will simply start reporting some well established - whenever known - criteria and methodology within each cluster of semantic web applications (§ 3).

However, at this stage (D21) we would like to clearly make a distinction between:

1. Internal properties of Knowledge Based system (e.g. syntactic anomalies : tautologies, cycles, …) as usually performed by the verification community
2. Service and Business level evaluation (e.g. utility, return-on-investment, economic suitability, …)

We claim that D21 must only address the level of Service and Business (level 2).
Few authors have tried to face the problem and so, few reported first proposals and practical conclusions are available (See the selected references in § 2.3)
The objectives of the next report (D22) will be to progress on that matter, in order to outline the generic properties of an evaluation framework for Service and Business assessment.

The major classical objectives of the evaluation are:

- to verify, demonstrate and evaluate the technical and economic suitability of the system to commercial applications,
- to evaluate the Ontology-based application results and provide recommendations of refinements and enhancements
- to define rough characteristics of business models and fields for which they could be useful.

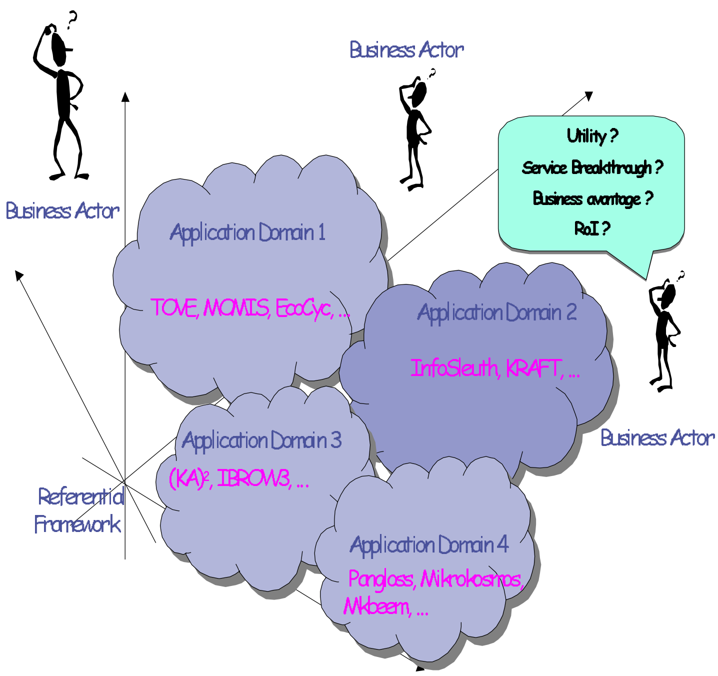


Fig. 1 Facing Evaluation of Ontology-Based Systems

[3] This chapter is simply introduced in that first release of the D21 series

The main purpose of the evaluation remains to carefully precise the real innovation of the "new" technology, application or service, even different from that assumed by the proposal or initially sought :

- ✓ **what was the state of the art improved in?** It may lie in several places : for instance, in a new ordering of existing functionalities, bringing thereby new or stronger services (as for robustness / relevance / quickness / ease of use / …) ; or in the technical field, opening possibilities of radically new services ; or in the strength of the gathered system developer team, eager to develop suitable solutions to some of their questions, and willing or not to offer them on the market ; … Most of the time, it is of course a mix of these, the evaluation purpose being then to split the different levels and precise each of them,
- ✓ **what is this innovation worth?** What are its worth and the worth of its uses in its supposed different application fields (they may be threefold : breakthrough, reinforcing or economical), and the value proposition it can support? How long could it give a differentiation advantage to its users and owners? What efforts and means would be necessary to fix and feed this advantage? Is there a sufficient plinth for so doing, as for available skills, technical documentation, …
- ✓ **what are the potential competitors, whether existing or to emerge?** In particular, what other technical means, possibly simpler or more efficient, may emerge within some time? How follow their emerging?
- ✓ **what marketing schemes seem the most relevant, throughout months**: first targeted fields, first envisaged restricted uses of the product, kinds of partnership to seek, … ? What particular business models could be derived from these schemes?

The point is to "file down" the system results, to reduce the system to its very gist, out of all its clothing, essential though when it comes to judge its usability (evaluating the product in a peculiar use, that of the retained business cases, while perceiving all its generic interest) ; this cannot be done by the contender team on its own, even if it is the only one able to lead this evaluation.

The team in charge of evaluation purpose is as a consultancy: commissioned by the project leader, it must study round the product, make up its mind its own way through finding support even technical outside the contender team, and stand up for its point of view in front of its client ; this latter remains the sole judge of the relevance of the conclusions and master of their publication – without twisting them.

The different pieces of advice, opinions, feelings will usually be gathered :

- ✓ **that of each developer (or developing team):** among other questions, what are they proud of, how do they express the value proposition of the developed product, what did they learn, what original knowing did they summon up, what improvement could be brought, what are the remaining difficulties, what were the initial objectives modified in, who masters the whole system?
- ✓ **that of each test user:** among other points, what is the perceived innovation and what is its supposed importance (worth of its uses), in what field should it be the most immediately interesting, what business model appear to be the most obvious as for them?
- ✓ **that of the first potential clients**: the commercial or marketing teams of the partners associated to the development are questioned on the same topics than test users, plus improvements and refinements necessary before adding the product into their existing commercial offer ; if allowed by the project achievement, it will also include testing of packaged pre-commercial proposals (including first version of commercial leaflet, issues of the possible commercial proposals, proposed model for sharing the generated value),
- ✓ **that of external technical consultants:** personnel from web agencies or technical companies, specialists of the field concerned by the innovation, they may have been met along the project in professional exhibitions ; they are to authenticate the technical content of the innovation, discuss the conditions of its development (including giving pieces of advice on the necessary characteristics of the needed personnel), and be its first supporters by its future clients and commercial partners.

Depending on the system achievement, this advice could be that of one big consultancy, specialised in technology management and evaluation (eg, Gartner Group, Forrester Research, IDC, Hurwitz Group, Standish Group).

On top of answering each main questions as to the real innovative content of the product, the evaluation team will give advice on the content of each of the 20 documents or media necessary to the product bringing into the market :

1. partners and partnership presentation
2. technology presentation
3. datasheets
4. technical white paper
5. return on investment
6. testimonials
7. competition analysis
8. evaluation package
9. demo application
10. rolling demonstration
11. installation guide
12. trouble shooting
13. getting started or methodology
14. user’s manual
15. courseware
16. qualification questions
17. price list + configuration questions
18. frequently asked questions
19. proof of concept
20. executive summary

2.3 References

[Grüninger et al.] Michael Grüninger, Mark S. Fox, **"Methodology for the Design and Evaluation of Ontologies",** Proceedings of IJCAI'95, Workshop on Basic Ontological Issues in Knowledge Sharing, April 13, 1995

[Cohen, 1995] **Empirical Methods for Artificial Intelligence**, MIT Press

[Menzies, 1997] T. Menzies, "**Evaluation Issues for Problem Solving Methods**" , Banff KA workshop 1998

[Menzies, 1998] T. Menzies, "**Evaluation Issues with Critical Success Metrics**", Proceedings of KAW’98

[Nick 1999], Nick M., Althoff K., Tauz C. , **“Facilitating the Practical Evaluation of Knowledge-Based systems and Organisational Memories using Goal-Questions-Metric Technique”,**.Proceedings of KAW99

[Jasper et al. 1999] R. Jasper and M. Uschold, " **A framework for Understanding and Classifying Ontology Applications** ", Proceedings of KAW’99

[Benjamin et al. 1999], *Asuncion Gomez Perez, V. Richard Benjamins*, **"Overview of Knowledge, Sharing and Reuse Components: Ontologies and Problem-Solving Methods**". In Proceedings of the IJCAI-99 workshop on Ontologies and Problem-Solving Methods (KRR5) Stockholm, Sweden, August 2, 1999. pp. 1.1--1.15

[Sure et al. 1999] York Sure, Rudi Studer, "**On-To-Knowledge Methodology - Employed and Evaluated Version**", On-To-Knowledge Project IST-1999-10132

[Julieanne et al. 2000 a] *Julieanne van Zyl, Dan Corbett,* **"Population of a Framework for Understanding and Classifying Ontology Applications"** *ECAI'2OOO*

[Julieanne et al. 2000 b] *Julieanne van Zyl, Dan Corbett,* **"A Framework for Comparing the use of a Linguistic Ontology in an Application",** *ECAI'2000*

[Rausher 1999] J. Rauscher, S. Marc, "**A la conquête de la Silicon Valley**" , Editions d'Organisation, Paris, ISBN: 2-7081-2282-7

3 Most Representative Applications

3.1 Introduction

In the field of Knowledge based Systems, we have identified so far the following key clusters of applications:

- Corporate Intranet and Knowledge Management
- E-Commerce (B2B, B2C)
- Information Retrieval
- Portals and Web communities

In each of those clusters, we provide

- An overall functional service definition of the application cluster
- An illustration through Main Use-Case[4]
- The selected most representative applications (2) or "killer application"
- References for further information

[4] *Use-Case diagrams* are notations defined in UML, the “Unified Modelling Language” (largely used in Industrial Software Engineering, and considered for extensions for the Knowledge Engineering emerging industry), which is basically a set of useful representations. Typically they are used during the early phases of system engineering and show an abstraction level of the services offered by the system. Use-Case put in the scene *actors* who stand outside the system and who interact with the system through services offered by the system.

3.2 Corporate Intranet and Knowledge Management

3.2.1 Service definition

Nowadays, knowledge is one of the most crucial success factors for enterprises. Therefore, Knowledge Management (KM) has been identified as a strategically important means for enterprises. Clearly, KM is an interdisciplinary task, including human resource management, enterprise organization and culture as well as IT technology. However, there is a wide-spread consensus that IT technology plays an important role as an enabler for many aspects of a KM solution.

In the past, IT technology for knowledge management has focused on the management of knowledge containers using text documents as the main repository and source of knowledge. In the future, Semantic Web technology, especially ontologies and machine-processable relational metadata, pave the way to KM solutions that are based on semantically related knowledge pieces of different granularity: Ontologies define a shared conceptualization of the application domain at hand and provide the basis for defining metadata that have a precisely defined semantics and are therefore machine-processable. Although first KM approaches and solutions have shown the benefits of ontologies and related methods there still exists a large number of open research issues that have to be addressed in order to make Semantic Web technologies a complete success for KM solutions:

Industrial KM applications have to avoid any kind of overhead as far as possible. Therefore, a **seamless integration** of knowledge creation, e.g. content and metadata specification, and knowledge access, e.g. querying or browsing, into the working environment is required. Strategies and methods are needed that support the creation of knowledge as side-effects of activities that are carried out anyway. This requires means for **emergent semantics**, e.g. through ontology learning, which reduces the overhead of building-up and maintaining ontologies.

Access to as well as presentation of knowledge has to be **context-dependent**. Since the context is set-up by the current business task and thus by the business process being handled, a tight integration of business process management and knowledge management is required. KM approaches being able to manage knowledge pieces provide a promising starting point for smart push services that will proactively deliver relevant knowledge for carrying out the task at hand.

Contextualization has to be supplemented by **personalization**. Taking into account the experience of the user and his or her personal needs is a prerequisite, on the one hand, for avoiding information overload and, on the other hand, for delivering knowledge on the right level of granularity.

The development of knowledge portals serving the needs of companies or communities is still a more or less manual process. Ontologies and related metadata provide a promising conceptual basis for generating parts of such knowledge portals. Obviously, conceptual models of the domain, the users and the tasks are needed among others. **Generation of knowledge portals** has to be supplemented with the (semi-)automatic evolution of portals. Since business environments and strategies change rather rapidly, KM portals have to be kept up-to-date in this fast changing environment. Evolution of portals also includes the aspect of ‘forgetting’ outdated knowledge.

KM solutions will be based on a combination of intranet-based functionalities and mobile functionalities in the very near future. Semantic Web technologies are a promising approach to meet the needs of the mobile environments, like e.g. location-aware personalization and adaptation of the presentation to the specfic needs of mobile devices, i.e. the presentation of the required information at an appropriate level of granularity. In essence, employees should have access to the KM application **anywhere** and **anytime**.

Peer-to-Peer computing combined with Semantic Web technology will be an interesting path to get rid of the more centralized KM solutions that are currently implied by ontology-based solutions. P2P scenarios open up the way to derive consensual conceptualizations among employees within an enterprise in a bottom-up manner.

Virtual organizations become more and more important in business scenarios that are characterized by decentralization and globalization. Obviously, semantic interoperability between different knowledge sources as well as trust is a must in inter-organizational KM applications.

The integration of KM applications, e.g. skill management, with **eLearning** is an important field enabling a lot of synergy between these two areas. KM solutions and eLearning have to be integrated from both an organizational and an IT point of view. Clearly, interoperability and/or integration of (metadata) standards are needed to realize such an integration.

Knowledge Management is obviously a very promising area for exploiting Semantic Web technology. Document-based KM solutions have already reached their limits, whereas semantic technologies open the way to meet the KM requirements of the future.

3.2.2 Use Case and Needs

The scenario shown in Figure 1, builds on the distinction between *know-edge process* (handling knowledge items) and *knowledge metaprocess* (introducing and maintaining KM systems).
Ontologies constitute the glue that binds knowl-edge sub-processes together. Ontologies open the way to move from a document-oriented view of KM to a content-oriented view, where knowledge items are interlinked, combined, and used.
The scenario shows that you can clearly identify and handle different sub-processes that drive the development and use of KM applications. You support these sub-processes by appropriate tools that are tied together by the ontology infrastructure.

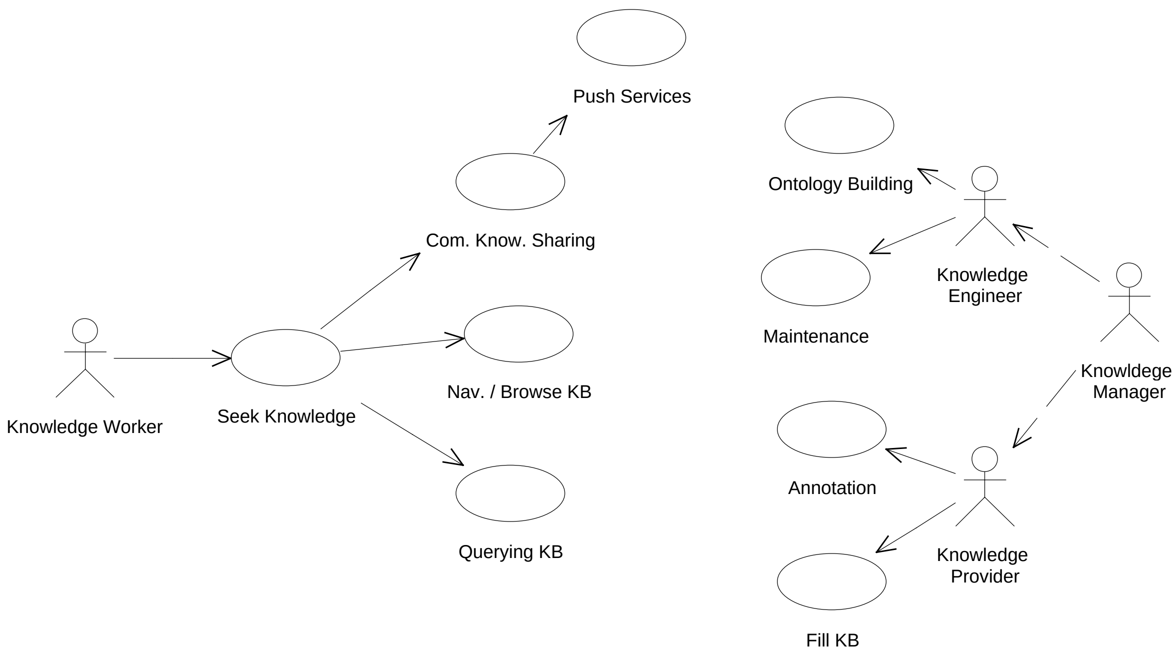


Figure 3.1 KM Use Case Two orthogonal processes with feedback loops illustrate an approach to knowledge management.

3.2.3 Most Representative Applications

3.2.3.1 Business-to-employee (B2E)

One of the key applications or may-be **Killer Application** of corporate Intranet is its usage in the **Mobile e-business** has the potential to deliver business-to-employee interactions capable of creating significant competitive edge. The real-life advantages of mobile access to business critical applications require that all organizations, including technology builders, have strategies in place for interacting with their own employees.

[Dyson, IBM 2001] The **business-to-employee (B2E)** opportunity will be the starting place for the real, extended application of mobile e-business. Mobile e-business is a natural progression of e-business – delivering applications and services to employees on the move. Importantly, mobile e-business is not only for companies selling consumer leisure gadgets, but also for organizations seeking to achieve added productivity, increased profitability, improved competitiveness and optimal return on investment.
The key to making the most of mobile e-business for B2E interactions is connecting the infrastructure and mobile device to line-of-business systems – Intranet collaboration, client billing information, client databases, production lines, e-mail, calendars and customer relationship management applications. In short, providing personalized and specific ways to take advantage of company information and resources.

3.2.3.2 OntoBroker

OntoBroker is one of the first implemented tools powered by the use of ontology and Semantic Web technology. It contains three core elements: a query interface for formulating queries, an inference engine used to derive answers, and a webcrawler used to collect the required knowledge from the web. It targets the support for providing a service that can be used more generally for the purpose of knowledge management and for integrating knowledge-based reasoning and semiformal representation of documents. The query formalism is oriented toward a frame-based representation of ontologies that defines the notion of instances, classes, attributes and values (D. Fensel, et al., 1998). OntoBroker has been successfully used in several user case scenarios:

- Semantic Community Web Portals: A Portal for the Knowledge Acquisition Community
- Knowledge Annotation Initiative of the Knowledge Acquisition Community (KA)2
- ProPer: Human Resource Knowledge Management

URL: http://ontobroker.aifb.uni-karlsruhe.de/index_ob.html

3.2.3.3 OntoKnowledge

Content-driven Knowledge-Management through Evolving Ontologies

Efficient knowledge management has been identified as key to maintaining the competitiveness of organizations. Traditional knowledge management is now facing new problems triggered by the web, to name but a few, information overload, inefficient keyword searching, heterogeneous information integration, geographical-distributed intranet problem and so on. These problems will be tackled by the modern technology – called the Semantic Web Technology (Fensel, 2001 & to appear). On-To-Knowledge[5] (OTK) project is an important player devoting itself to finding the content-driven knowledge management solutions through evolving ontologies. It employs the power of the Semantic Web Technology to facilitate knowledge management.

URL: http://www.ontoknowledge.org

[5] www.ontoknowledge.org

tool structure

On-To-Knowledge supports efficient and effective knowledge management by providing a tool environment powered by the Semantic Web Technology. It focuses on *acquiring, maintaining, and accessing* weakly structured information sources:

- *Acquiring*: Text mining and extraction techniques are applied to extract semantic information from textual information (i.e., to acquire information). Tool support includes ontology extraction from text (OntoExtract and OntoWrapper)
- *Maintaining:* RDF, XML and DAML+OIL are used for describing syntax and semantics of semi-structured information sources. Tool support includes ontology editor (OntoEditor), and ontology storage and retrieval (Sesame, RDF-Ferret), so as to enable automatic maintenance and view definitions on these knowledge.
- *Accessing*: Push-services and agent technology support users in accessing the information. Tool support includes ontology-based information navigation and querying (RDF-Ferret), and ontology-based visualization of information (Spectacle).

In a nutshell, the complete layered tool environment of the OnToKnowledge is (see Figure below): OntoExtract and OntoWrapper extract unstructured and structured textual information sources from specified domains at the Internet or Intranet. The extracted information is pumped into the RDF-DB (Sesame), where it can be edited with the OntoEdit tool. Finally, the RQL (RDF querying language) reasoning engine allows for the querying of this database and delivers results to a user through a smart user- RDF-Ferret and visualized by Spectacle.

Real-life Applications

Three real-life applications have been conduct during the course of the OTK project to fulfil two requirements: - to identify the real-life requirement for the design of the tools and another way around - to secure the usability of the tools for tackling the real-life problems.

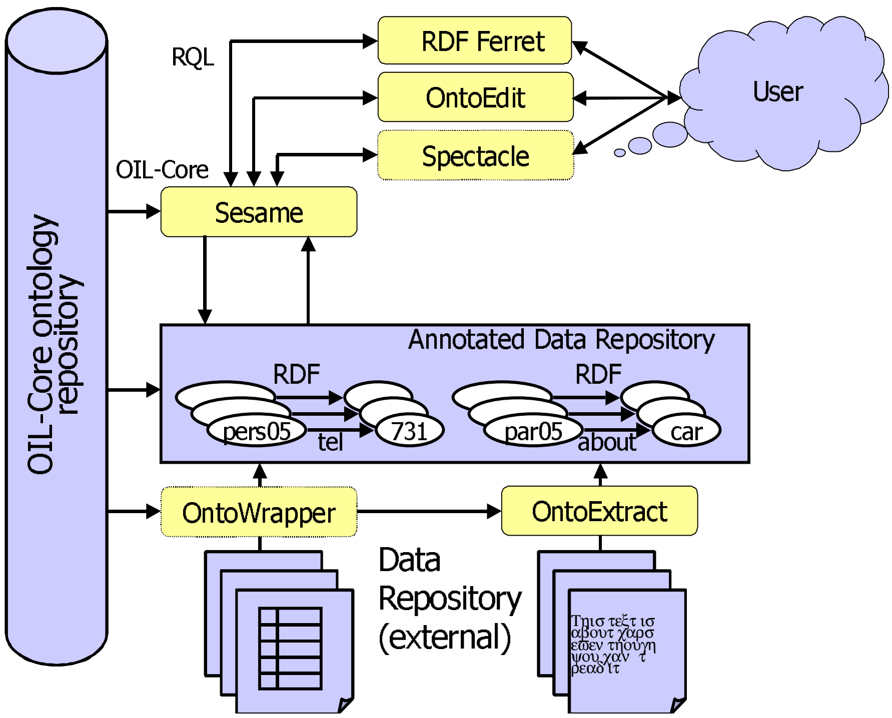


Figure 3.2 OnTo Knowledge The layered tool environment of the OnToKnowledge

BritishTelecom call center: Call center is the platform for companies to communicate with their customers and its market is growing 20% a year, with millions being spent on improving customer relationships. Current call center technology lacks of the support of the operator in solving incoming requests. The investment in call center technology can offer great rewards, including better customer service, lower overheads, lower operational costs, and increased staff profitability. In the BT case study, a system for supporting intranet-based virtual communities of practice is being developed, allowing the automatic sharing of information. The system, *OntoShare*, allows the storage of best practice information in an ontology and the automatic dissemination of new best practice information to relevant call center agents. In addition, call center agents can browse or search the ontology to find information of most relevance to the problem they are dealing with at any given time. The ontology helps to orientate new agents and acts as a store for key leanings and best practices accumulated through experience. It provides a sharable structure for the knowledge base, and a common language for communication between call center agents.

Swiss Life Applications: Two of the case studies were carried out by Swiss Life. One of it approached the problem of finding relevant information in a very large document about the International Accounting Standard (IAS) on the extranet. With the help of the ontology extraction tool OntoExtract an ontology was automatically learned from the document, which significantly supports a user in reformulating an initial query when it did not deliver the intended results. The second case study of Swiss Life is a skill management application that uses manually constructed ontologies about skills, job functions, and education. This enables an employee to create in a simple way a personal home page on the company's intranet that includes information about personal skills, job functions, and education. Using the ontology allows a comparison of skills descriptions among employees, and makes it possible to automatically extend a query with more general, or more specialized, or semantically associated concepts.

Enersearch Applications: The case study done by EnerSearch AB focuses on validating the industrial value of the project results with respect to the needs of a virtual organization. The main difficulty with the current web site is that it is rather hard to find information on certain topics because the current search engine supports free text search rather than content-based information retrieval. To improve, the whole web site is annotated by concepts from an ontology which was developed using a semi-automatic extraction from the documents on the existing EnerSearch web site. The RDFFerret search engine is used to extend the free text search to a search of the annotations as well. Alternatively, with the Spectacle tool a user is able to get a search result arranged into a topic hierarchy which can then be browsed, allowing the user to look for the required information in a more explorative way.

3.2.3.4 CoMMA (Corporate Memory Management through Agents).

Objectives

The COMMA project aims at implementing a corporate memory management framework based on agent technology, that will capture, store and diffuse embedded knowledge of many types, in interactive sessions to employees.

The CoMMA project will address particularly the following issues:

Enhancing the insertion of new employees in the company by capturing experience and know-how from elder employees. The facility and the rapidity to integrate new employees are a major issue in a competitive market.

Performing processes that detect, identify and interpret technology movements and interactions for matching technology evolutions with market opportunities to diffuse among employees innovative ideas related to technology monitoring activities.

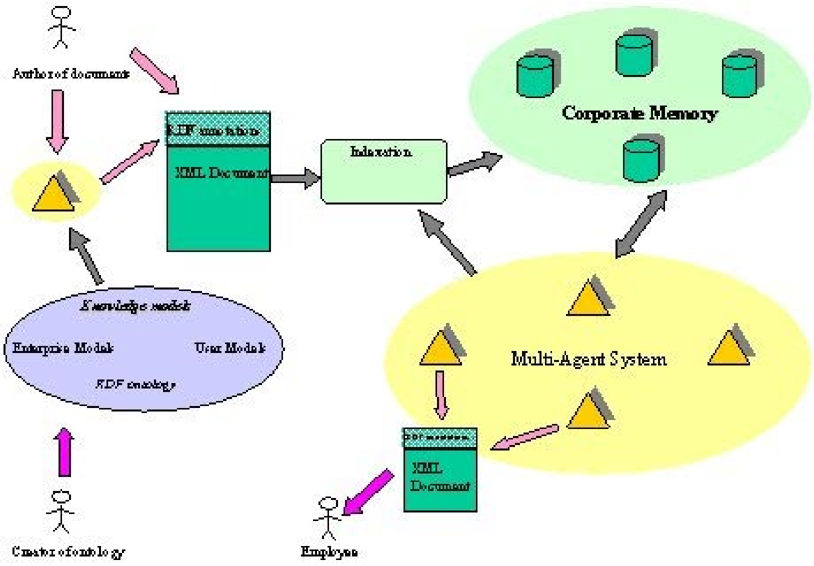


Figure 3.3 : COMMA Overall Architecture

COMMA project will design this corporate memory by merging most innovative technologies:

- Multi agent system,
- XML, RDF format,
- machine learning techniques

COMMA project will provide new services, products and tools for knowledge management as dedicated agents able to achieve specific tasks for information retrieval

URL: http://www.si.fr.atosorigin.com/sophia/comma/Htm/HomePage.htm

3.2.4 References

[Staab et al., 2001], Steffen Staab, Rudi Studer, Hans-Peter Schnurr, York Sure, **"Knowledge Management : Knowledge Processes and Ontologies"**; IEEE Intelligent Systems Journal, Jan. Feb. 2001.

[Dieng et al., 1999], R. Dieng, O. Corby, A. Giboin, and M. Ribiere. **Methods and tools for corporate knowledge management**. Int.Journal of Human-Computer Studies, 51(3):567–598, 1999.

[Kuehn et al., 1997], O. Kuehn and A. Abecker. **Corporate memories for knowledge memories in industrial practice: Prospects and challenges**. *Journal of Universal Computer Science*, 3(8), August 1997.

[Schreiber et al., 1999], G. Schreiber, H. Akkermans, A. Anjewierden, R. de Hoog, N. Shadbolt, W. Van de Velde, and B. Wielinga. **Knowledge Engineering and Management — The CommonKADS Methodology**. The MIT Press, Cambridge, Massachusetts; London, England, 1999.

[Dyson S. IBM, 2001], S. Dyson, IBM UK, « **B2E Mobile e-Business : Driver, Passenger or Spectator ?** », in [E-work and E-commerce 2001].

[D. fensel, 2001] D. Fensel: "**Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce**". Springer-Verlag, Berlin, 2001.

[D. Fensel et al. 2002] D. Fensel, J. Hendler, H. Lieberman, and W. Wahlster (eds.): "**Semantic Web Technology**", MIT Press, Boston, to appear.

[D. Fensel, 1998], D. Fensel, S. Decker, M. Erdmann, and R. Studer: "**Ontobroker in a Nutshell**" (short paper). In C. Nikolaou et al. (eds.), Research and Advanced Technology for Digital Libraries, *Lecture Notes in Computer Science*, LNCS 1513, Springer-Verlag Berlin, 1998.

[O'Leary, et al. 2001] Dan E. O'Leary and Rudi Studer (eds.): "**Knowledge Management**", IEEE Intelligent Systems 16 (1), January/February 2001 (special issue).

[Staab S. et al. 2001] Steffen Staab and Alexander Maedche, "**Knowledge Portals — Ontologies at Work**". *AI Magazine* 21(2), Summer 2001.

IST Knowledge Management portal , http://www.knowledgeboard.com/

3.3 E-Commerce

3.3.1 Service definition

Electronic Commerce is based on the exchange of information between involved stakeholders using a telecommunication infrastructure. There are two main scenarios: Business-to-Customer (B2C) and Business-to-Business (B2B).

B2C applications enable service providers to propagate their offers, and customers to find offers which match their demands. By providing a single access to a large collection of frequently updated offers and customers, an electronic marketplace can match the demand and supply processes within a commercial mediation environment.

B2B applications have a long history of using electronic messaging to exchange information relating to services previously agreed between two or more businesses. Early plain-text telex communication systems were followed by electronic data interchange (EDI) systems based on terse, highly codified, well structured, messages. Recent developments have been based on the use of less highly codified messages that have been structured using the eXtensible Markup Language (XML).
A new generation of B2B systems is being developed under the ebXML (electronic business in XML) label. These will use classification schemes to identify the context in which messages have been, or should be, exchanged. They will also introduce new techniques for the formal recording of business processes, and for the linking of business processes through the exchange of well structured business messages. ebXML will also develop techniques that will allow businesses to identify new suppliers through the use of registries that allow users to identify which services a supplier can offer.

The coding systems used in EDI systems are often examples of limited scope, language independent, mini-ontologies that were developed in the days when decimalised hierarchical classification systems were the most sophisticated form of ontology. There is a strong case for the redesign of many of these classification schemes based on current best practice for ontology development. ebXML needs to include well managed multilingual ontologies that can be used to help users to match needs expressed in their own language with those expressed in the service providers language(s).
Within Europe many of the needs of B2C applications match those of B2B applications. Customers need to use their own language to specify their requirements. These need to be matched with services provided by businesses, which may be defined in languages other than those of the customer. Businesses may or may not provide multilingual catalogues. Even where multilingual catalogues are supplied they may not cover all European languages. For a single market to truly exist within Europe it must be possible for customers to be able to request product and sales term information in their own language, possibly through the use of on-line translation services. It is anticipated that services providing multilingual searching of sets of catalogues will act as an intermediary between businesses and their potential customers.

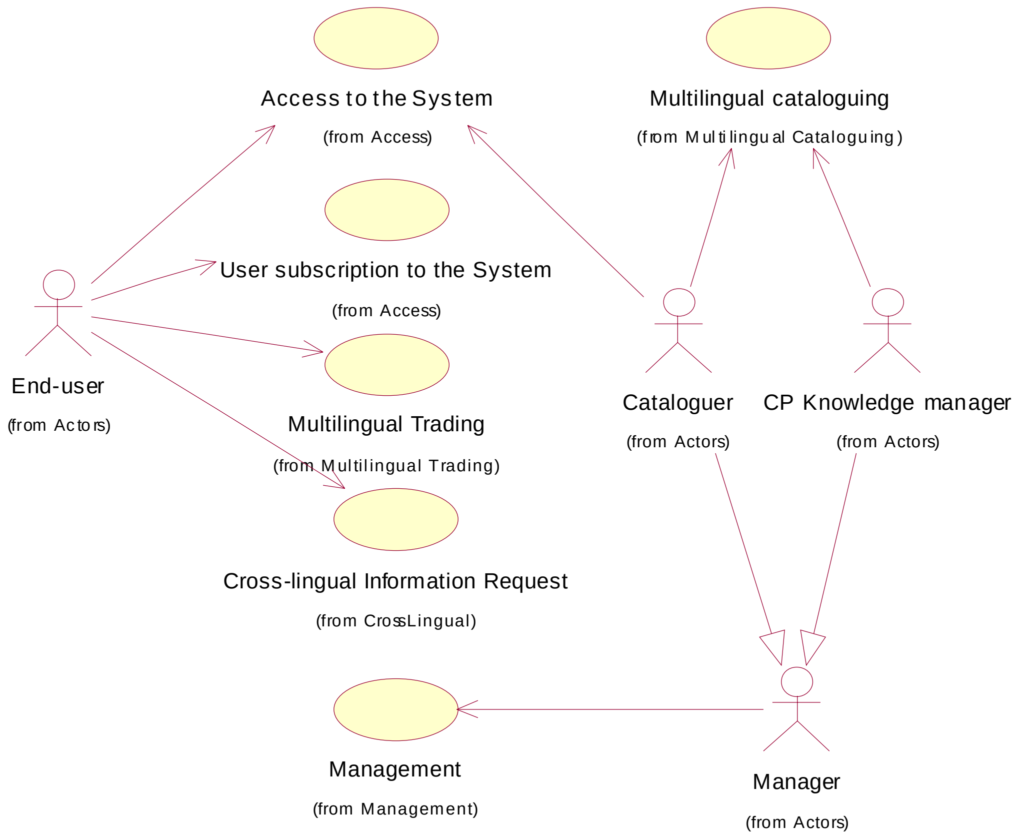
3.3.2 Use Cases and Needs in B2C

For B2C transactions customers need access to electronic systems that:

- Provide a user interface in their own language
- Provide access to details of products from a range of different firms that create products of the type required. For adequate comparison of products their descriptions will need to be “standardized” so that they include comparable qualities.
- Allow users to request searches of multiple supplier catalogues using natural language terms so that the system can help them to identify which suppliers supply relevant products. For adequate response

to search requests search engines will need to either use "standardized" sets of product description properties or use ontologies that provide facilities for identifying equivalences between terms.

- Provide users with a means of paying for the product that is acceptable and accessible to them and which the supplier can also accept. (This may require the services of one or more third party service suppliers, either in the form of credit card or banking services, or of "mall operators" willing to act as debt intermediaries.)
- Provide means for the timely and properly managed delivery of products from supplier to customer, including facilities for the automatic identification of most appropriate supply point for the requested delivery point.



3.3.6 Use Cases and Needs for B2B

The following needs exist:

- Trading partners will need to be able agree on the set of terminology that will be used to describe the products to be traded with respect to their own domains. (For example, a Dutch adhesive manufacturer supplying wood glue to an Italian furniture maker will need to use terms and property sets that are relevant to Italian furniture makers to describe their products.)
- Need to assign the labels used to identify specific fields within a message in a form that is dependent on locale and domain specific terminology (e.g. the terms used to describe transportation within Finland, or the terms used to describe viniculture in Hungary)

- Need to be able to express the options provided in choice lists as locale-specific terms those creating or receiving messages can understand in their own language
- Need to use terms and conditions that have previously been translated into the language of the business accepting the terms (which could be the supplier or the customer in the B2B transaction)
- Need to adjust the legal and financial constraints on transactions to take account of the source of supply and the point of delivery. (For example, a company in Greece could ask a company in Portugal to ship goods from its Swedish warehouse to a delivery point in Denmark. Whilst the contract may be between Greek and Portuguese company, in terms of the laws of one of these countries, the financial information used for customs purposes must be expressed in terms acceptable to the Danish and Swedish authorities, in the currencies they are currently using.)

3.3.7 Representative Applications

3.3.7.1 IST-MKBEEM

The MKBEEM platform focuses on adding multilinguality to the following stages of the information cycle for multilingual B2C portal services: products or services content and catalogue semi-automated maintenance; automated translation and interpretation of natural language user requests, and natural dialogue interactivity and usability of the service making use of combined navigation and natural language inputs.
The main overall goals of MKBEEM are to: develop intelligent knowledge-based multilingual key components (NLP and KRR) for applications in a multilingual electronic commerce platform; validate and assess the prototypes on a pan-European scale (France and Finland) with three basic languages (Finnish, English and French) and two optional languages.

3.3.7.2 CEN/ISSS MULECO

The Multilingual Upper-Level Electronic Commerce Ontology (MULECO) currently being proposed by the CEN/ISSS Electronic Commerce Workshop is designed to provide a mechanism whereby relationships between the high-level terms in business ontologies can be inter-connected. Recognizing that most existing ontologies are domain and language specific, and that there is a need to be able to relate terms in one ontology with those in another ontology as part of the semantic translation that is needed to inter-relate applications (as defined within the E-Commerce Integration Meta-Framework (ECIMF) being defined by the CEN/ISSS EC Workshop), the MULECO team are hoping to define techniques that will allow multilingual querying of ontologies based on the relationships between the local ontologies and a set of well defined business and business process classification schemes.

3.3.7.3 IST-SmartEC

The goal of Smart-EC is to build the e-Trade platform on which the buyer will be able to purchase a Global Service (of the user’s demand requires several providers and involves more than one actors) as a single service. During the first stage of the negotiation Smart-EC will help the user defining his/her demand and will propose one or more global and convenient solutions. After the user agreement, Smart-EC will be in charge of confirming the transaction to the respective providers and to distribute the payment to the respective third parties.

The Global Service will be broken in more atomic services according to the knowledge modelled by Ontologies. The Ontology rules will modelled the generic / specific relations between services, the time and coherence constraints between services. Rules will also specify the commercial behaviour of services such as the necessity to check availability on-line.

3.3.8 References

[MKBEEM 2000]
URL: http://www.mkbeem.com/

[MULECO 2001]
URL: http://www.cenorm.be/isss/Workshop/ec/MULECO/Documents/Muleco_Documents.htm

[SmartEC, 2000]
URL: http://www.telecom.ntua.gr/smartec/

[E-work and E-commerce 2001] "**Novel solutions and practices for a global networked economy**", e.2001, 17-20 October 2001, Venice, Italy. Edited by Brian Stanford-Smith and Enrica Chiozza, ISBN: 1 58603 205 4.
URL: http://www.ebew.net/programme.htm

[eWork 2001], "Status Report on New Ways to Work in the Knowledge Economy", IST document
URL: http://www.eto.org.uk

3.4 Information Retrieval

3.4.1 Service definition

In an information retrieval (IR) application, ontologies are used to guide the search so that the system may return more relevant results. The assumption in this class of application is that the ontology will allow the IR system a better representation ("understanding") of the concepts being searched and thus make possible an improvement of its performance from what is presently the case.

The problems of IR are well known to the research and user communities. Amongst the most widely recognized ones are the so-called *missed positives* and *false positives*. In the first case the system fails to retrieve relevant answers to the query whereas in the second case the system retrieves answers that are irrelevant to the query. Throughout the years a number of mechanisms of various levels of sophistication have been devised for ranking the results produced from the search. These mechanisms range from human based indexing, to statistical measures (word frequency analysis) and more recently to Internet-related measures such as page popularity and number of incoming links.

While the human-based method might be considered the best for most cases, the rapid increase in the amounts of information presently being generated makes this an increasingly untenable proposition. The use of ontology-enhanced search and retrieval promises to address this issue since it attempts to replicate the level of quality of human-based representation of information and concepts while still being able to handle vast amounts of data.

However, even such an approach risks failing if it does take into account generic search issues that have not been tackled in the past, as well as some issues that are specific to search through the Internet. The main issues are:

- *Context*: context is the conceptual framework that determines how relevant a query result is.
- *Information quality*: This refers to issues such as how up to date the information is, the existence of different versions of the same information as well as the existence of contradictory information
- *User search mode*: the goal of a search may often differ ranging from browsing, to discovery, to focused search. In each case the user will be expecting increasingly relevant results, which has a direct bearing on the search and ranking mechanisms employed.

3.4.2 Use Cases and Needs

The following figure illustrates an application scenario that addresses the issues mentioned above. Ontology creation and maintenance are not indicated. On the other hand the architecture shown distinguishes between internal (proprietary) and external (non-proprietary) sources indicating the need for ontology alignment in the latter case.

Note also the central role of the ontology both in the 'query augmentation' phase and in the processing of the initial results. In the former case the ontology will form the basis for better understanding the context of the user's initial query taking into account the search mode employed.

The ontology is also used in the filtering, ranking and presentation of the results covering quality issues such as contradictions and related information (a different possible answer to the same query or an answer to a different but related query).

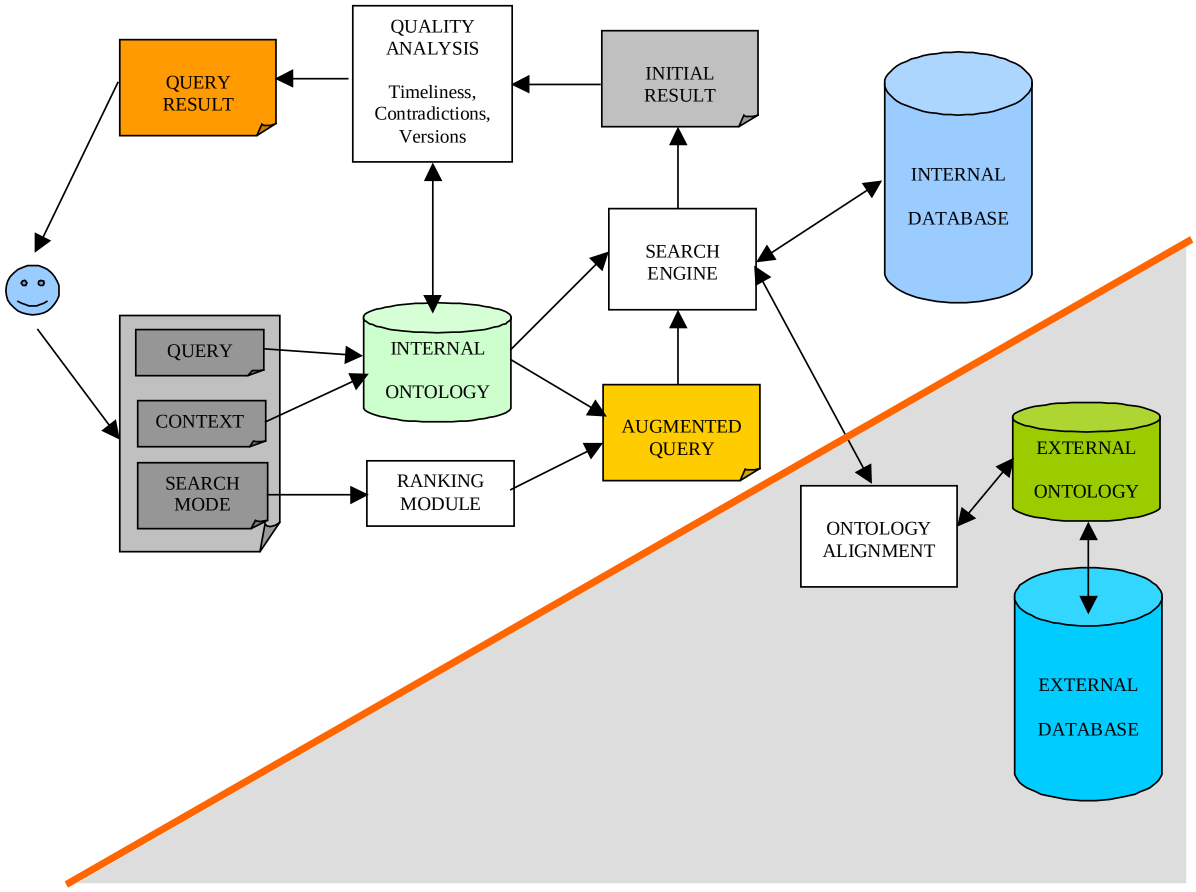


Figure 3.4. 1: Basic architecture for ontology based information retrieval applications

What to evaluate?
The usual criteria used in the IR community are mostly the following

- **Coverage of the collection**: extent to which the system is able to cover all material available (e.g. scaling up to the Web size)
- **Time for answering**: average interval between the time request is made and the time answer is given.
- **Presentation of the Results**
- **Effort** to obtain answers
- **Recall**: proportion of relevant material actually retrieved
- **Precision**: proportion of retrieved material that is actually relevant

Ontology-Based Information Retrieval Specific Evaluation Criteria

The criteria discussed above are 'generic' criteria in that they are applicable irrespective of their intended use of an ontology. The following is a list of criteria specific to information retrieval:

- What are the **improvements on the precision and recall** of the system?
- Does the ontology allow for the **representation of context of the Query formulation and answer**?

- What are the special features of the query language? Does it support the **augmenting of the user's query**?
- Is there a **ontology-based visualization** system available that **supports query formulation**?
- How are query results better ranked compared to traditional IR systems?
- Does the system allow for different kinds of search modes e.g. specific search, browsing, discovery?
- Does the system provide for **explanations** on Query answers or better on the no-answer cases?

3.4.3 Representative Applications

3.4.3.1 OntoSeek

For Yellow pages and products catalogues, structured content representations coupled with linguistic ontologies can increase both recall and precision of content-based-retrieval. The OntoSeek system adopts a language of limited expressiveness for content representation and uses a large ontology based on Wordnet for content matching.

Ontoseek combines an ontology-driven content-matching mechanism with a moderately expressive representation formalism. Differently from most of current systems, the user is not assumed to have familiarity with the vocabulary used for component encoding, but the system relies on a large linguistic ontology called Sensus to perform the match between queries and data. It assumes that the information encoding and retrieval processes will involve a degree of interactivity with a human user.

3.4.3.2 SARI - A System for Semantical Information Retriev al

The SARI (Software Agents for Retrieval of Information) system is intended to act as a broker between human users or other computerized systems (ie applications) needing information at one end, and heterogeneous information sources with different search engines at the other.

SARI's architecture reflects the system's role as a broker between its users and information sources. SARI's has agents of the following types:

- Application Agents represent the users (humans or other computerized systems) to the SARI system.
- Search Agents mediate information sources. They compile queries coming from Control Agents into the query languages of their information sources, and send the results back to the Control Agents
- Control Agents act as brokers in the SARI system. Each Control Agent receives agent messages containing information retrieval requests from Application Agents, decides to which Search Agents it forwards the requests, sends messages containing the retrieval requests to the appropriate Search Agents, receives messages containing search results from the Search Agents, combines them into information retrieval results, and sends the retrieval results back to the Application Agents
- Ontology Agent contains metadata in the form of ontologies that describe the conceptual structure of the information present in the information sources used by SARI.

In addition, there are also Content Provider Agents that represent content providers to the SARI system. Content providers are organizations or individuals who own one or more information sources that are accessible to the SARI system. Control Agents form the heart of SARI. They make their brokering decisions on the grounds of the user information lying in user profiles, and of the metadata about the information to be retrieved lying in ontologies. Control Agents can form federations with each other, as a rule, but there is just one Control Agent in the present pilot version of SARI.

The conceptual structure of the information contained in the information sources available to SARI is described by ontologies. The ontologies describing Web resources are specified as RDF schemas and descriptions for SARI. Ontologies can be graphically browsed in SARI.

In SARI the concepts of different ontologies are linked to each other by making use of the notions of viewpoint and bridge. The ontologies interlinked in such a way form the ontological structure that can be viewed from different perspectives. For example, there is a bridge between the concepts Commodity and Product which are respectively the root classes of the classifications under the foreign trade and manufacturing viewpoints.
Future goals with SARI include making the formation of bridges between the concepts of different ontologies semiautomatic, and also semiautomatic generation of RDF metadata from Web resources.

The SARI system is being worked out in Finland jointly by VTT Information Technology, Tampere University of Technology, and Tampere University. The project started in March 1996, and will continue until March 1999.

3.4.3.3 Medical Concept Mapper

The Medical Concept Mapper is a tool which combines the Unified Medical Languages System (UMLS) developed by the National Library of Medicine (NLM) with automated computational search space tools developed by the AI Lab at the University of Arizona. The Medical Concept Mapper consists of five components that combine human-created and computer-generated elements. The Arizona Noun Phraser extracts phrases from natural language user queries. WordNet and the UMLS Metathesaurus provide synonyms. The Arizona Concept Space generates conceptually related terms. Semantic relationships between queries and concepts are established using the UMLS Semantic Net.

Medical Concept Mapper is used as an aid for providing synonyms and semantically related concepts to improve searching. All terms are related to the user-query and fit into the query context. The system is an in-depth integration of manually created ontologies and computer generated tools, the intertwining of which allows a synergy to surface that surpasses the weaknesses and strengths of each tool when used on its own.

MCM has been tested on two expert groups: medical librarians and cancer researchers.
The test showed that the system can automatically double the number of useful search terms extracted from queries. It can also suggest related terms with high precision.

3.4.4 References

[Guarino et al., 1999], Nicola Guarino, Claudio Masolo, Guido Vetere**, "OntoSeek: Content-Based Access to the Web**", IEEE Intelligent Systems, 1999.
[Stuart Aitken 2000] Stuart Aitken, Sandy Reid, **"Evaluation of an Ontology-Based Information retrieval Tool"**, ECAI'00, Applications of Ontologies and Problem-Solving Methods.
[Fensel et al., 2000], Dieter Fensel, Stefan Decker, Michael Erdmann, Rudi Studer, "**Ontobroker: Ontology based Access to Distributed and Semi-Structured Information"**
[Qi Li et al 2001], Qi Li, Philip Shilane, Natalya Fridman Noy, M.A. Musen "**Ontology Acquisition from on-line Knowledge Sources**", AMIA Inc. pp. 497 - 501
[L. Crow 2001], L. Crow, N. Shadbolt "**Extracting focused knowledge from the semantic web**" Int. J. Human-Computer Studies 54, pp 155-184
[S. A. McIlraith 2001], S.A. McIlraith, T.C Son, H. Zeng "**Semantic Web Services**". IEEE Intelligent Systems, pp. 46-53

SARI : http://www.ercim.org/publication/Ercim_News/enw35/taveter.html

MCM: http://ai.bpa.arizona.edu/go/intranet/papers/Customizable-00.htm

3.5 Portals and Web communities

3.5.1 Service definition

The widely-agreed core idea of the Semantic Web is the delivery of data on a semantic basis. Intuitively the delivery of semantically processable data should help with establishing a higher quality of communication between the information provider and the consumer. The vision of the Semantic Web is closely related to ontologies as a sound semantic basis that is used to define the meaning of terms and hence to support intelligent providing and access to information for Web communities.

Navigating through Web portals which are based on a topic thesaurus, like http://dmoz.org or http://www.yahoo.com[6], is more or less equivalent to browsing a static hierarchy. Those with a richer semantic model, such as KA2Portal [AIF00] (http://ka2portal.aifb.uni-karlsruhe.de), offer simple navigation through a class hierarchy.

The **Ontobroke**r project [Ontobroker 1999] lays the technological foundations for the AIFB portal. On top of Ontobroker the portal has been built and organizational structures for developing and maintaining it have been established. The approach closest to Ontobroker is **SHOE** [SHOE 2000]. In SHOE, HTML pages are annotated via ontologies to support information retrieval based on semantic information. Besides the use of ontologies and the annotation of web pages the underlying philosophy of both systems differs significantly: SHOE uses description logic as its basic representation formalism, but it offers only very limited inferencing capabilities. Ontobroker relies on Frame-Logic and supports complex inferencing for query answering. Furthermore, the SHOE search tool does not provide means for a semantic ranking of query results. A more detailed comparison to other portal approaches and underlying methods may be found in [Staab S. et al. 2000].

A richer semantic model, such as **SEAL Portal** [Stojanovic et al. 2001] or C-Web [Saglio et al. 2002], offer navigation through a class hierarchy for dynamic exploration.
Such framework should help users to navigate through portals leading into very large resource spaces and to find quickly many resources but only those of interest for them.

3.5.2 Use Cases and Needs

The overall architecture and environment of a classical portal for a Web community is well illustrated by [Stojanovic et al. 2001] see Figure SEAL. The *backbone* of the system consists of the *knowledge warehouse i.e.* the ontology and knowledge base, and the *Ontobroker* system, *i.e.* the principal inferencing mechanism. The latter functions as a kind of middleware run-time system, possibly mediating between different information sources when the environment becomes more complex than it is now.
At the front end one may distinguish between three types of *agents*: *software agents*, *community users* and *general users*. All three of them communicate with the system through the *web server*.
The three different types of agents correspond to three primary modes of interaction with the system.

[6] In contrast to the SEAL approach Yahoo only utilizes a very light-weight ontology that solely consists of categories arranged in a hierarchical manner. Yahoo offers keyword search (local to a selected topic or global) in addition to hierarchical navigation, but is only able to retrieve complete documents, *i.e.* it is not able to answer queries concerning the contents of documents, not to mention to combine facts being found in different documents or to include facts that could be derived through ontological axioms.

First, remote applications (*e.g.* software agents) may process information stored in the portal. For this purpose, the *RDF generator* presents RDF facts through the web server. Software agents with *RDF crawlers* may collect the facts and, thus, have direct access to semantic knowledge stored at the web site.

Second, Community users and general users can access information contained at the web site. Two forms of accessing are supported: navigating through the portal by exploiting hyperlink structure of documents and searching for information by posting queries. The hyperlink structure is partially given by the portal builder, but it may be extended with the help of the *navigation* module. The navigation module exploits inferencing capabilities of the inference engine in order to construct conceptual hyperlink structures. Searching and querying is performed via the *query* module. In addition, the user can personalize the search interface using the *semantic personalization* preprocessing module and/or rank retrieved results according to semantic similarity (done by the post-processing module for *semantic ranking*). Queries also take advantage of the Ontobroker inferencing capabilities.

Third, only community users can provide data. Typical information they contribute includes personal data, information about research areas, publications, activities and other research information. For each type of information they may contribute there is (at least) one concept in the ontology. By retrieving parts of the ontology, the *template* module may semi-automatically produce suitable HTML forms for data input. The community users fill in these forms and the template modules stores the data in the knowledge warehouse.

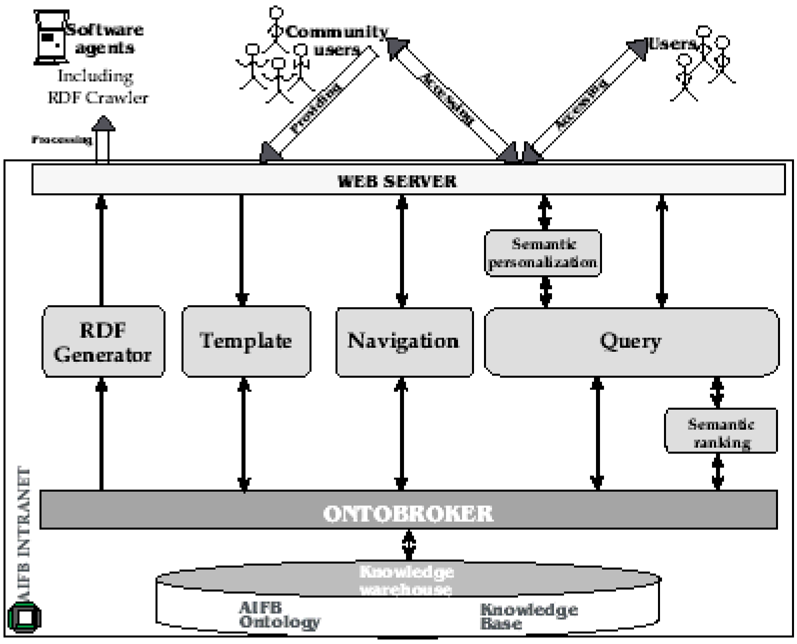


Figure 3.5.2 SEAL Overall architecture and Use Case of Classical Portal for Web Community

3.5.3 Representative Applications

SEAL - A Framework for developing Semantic Portals

The core idea of the Semantic Web is to make information accessible to human and software agents on a semantic basis. Hence, web sites may feed directly from the Semantic Web exploiting the underlying

structures for human and machine access. We have developed a generic approach for developing semantic portals, *viz.* SEAL (SEmantic portAL), that exploits semantics for providing and accessing information at a portal as well as constructing and
maintaining the portal.
In SEAL the focus is on semantics-based means that make semantic web sites accessible from the outside, i.e. semantics-based browsing, semantic querying, querying with semantic similarity, and machine access to semantic information. In particular, we focus on methods for acquiring and structuring community information as well as methods for sharing information.
As a case study we refer to the AIFB portal — a place that is increasingly driven by SemanticWeb technologies. We also discuss lessons learned from the ontology development of the AIFB portal.

C-Web : Community Webs
C_WEB is a collaborative effort, associating several research organizations, IT companies and "advanced users", aiming at designing a generic platform based on open standards and distributed as open source, and the related methodology and know-how, to support community-webs. C-Web consortium intends to validate both the software platform and the methodology through experiments carried-out with several user communities

3.5.4 References

[Stojanovic et al., 2001], N. Stojanovic, A. Maedche, S. Staab, R. Studer, and Y. Sure**. "SEAL— A Framework for Developing SEmantic PortALs"**. In *K-Cap 2001* **-** *First International Conference on Knowledge Capture, Oct. 21-23, 2001,* Victoria, B.C., Canada, 2001. to appear.
URL: http://ontobroker.semanticweb.org/ontos/aifb.html

[Saglio et al. 2002], J. M. Saglio, Tuang Anh Ta, "A Framework for Dynamic Exploration in Semantic Portals", KR2002 to appear
C-Web : Community Webs
URL: http://cweb.inria.fr/

[Staab S. et al. 2000] S. Staab, J. Angele, S. Decker, A. Hotho, A. Maedche, H-P. Schnurr, R. Studer, and Y. Sure. "**AI for the web — ontology-based community web portals**". In *AAAI 2000/IAAI 2000* **-** *Proceedings of the 17th National Conference on Artificial Intelligence and 12th Innovative Applications of Artificial Intelligence Conference*. AAAI Press/MIT Press, 2000.

[Ontobroker, 1999] S. Decker, M. Erdmann, D. Fensel, and R. Studer. "**Ontobroker: Ontology Based Access to Distributed and Semi-Structured Information**". In R. Meersman et al., editors, *Database Semantics: Semantic Issues in Multimedia Systems*, pages 351–369. Kluwer Academic Publisher, 1999.

[SHOE 2000] J. Heflin and J. Hendler. "**Searching the web with SHOE**". In *Artificial Intelligence for Web Search. Papers from the AAAI Workshop. WS-00-01*, pages 35–40. AAAI Press, 2000.

4 Evaluation of Ontology Based Applications

4.1 Evaluation Methodology

It is an important industrial need to deliver high-quality Knowledge-based systems. Evaluation is required to ensure this high quality and to guide the development and maintenance. Problems with existing evaluation studies [Menzies, 1998] show that there is a crucial need for a systematic methodology that helps in conducting "good measurements".

One known methodology refers to the Goal-Question-Metric (GQM) technique [Rombach, H.D. , 1991], which have been successfully applied to CBR-based Knowledge Management [Nick, 1999]. The GQM evaluation methodology includes the process, templates and guidelines for the application of GQM as proposed by [Briand et al., 1996]. The GQM methodology is an industrial-proven technique that has been successfully used in the Software industry.

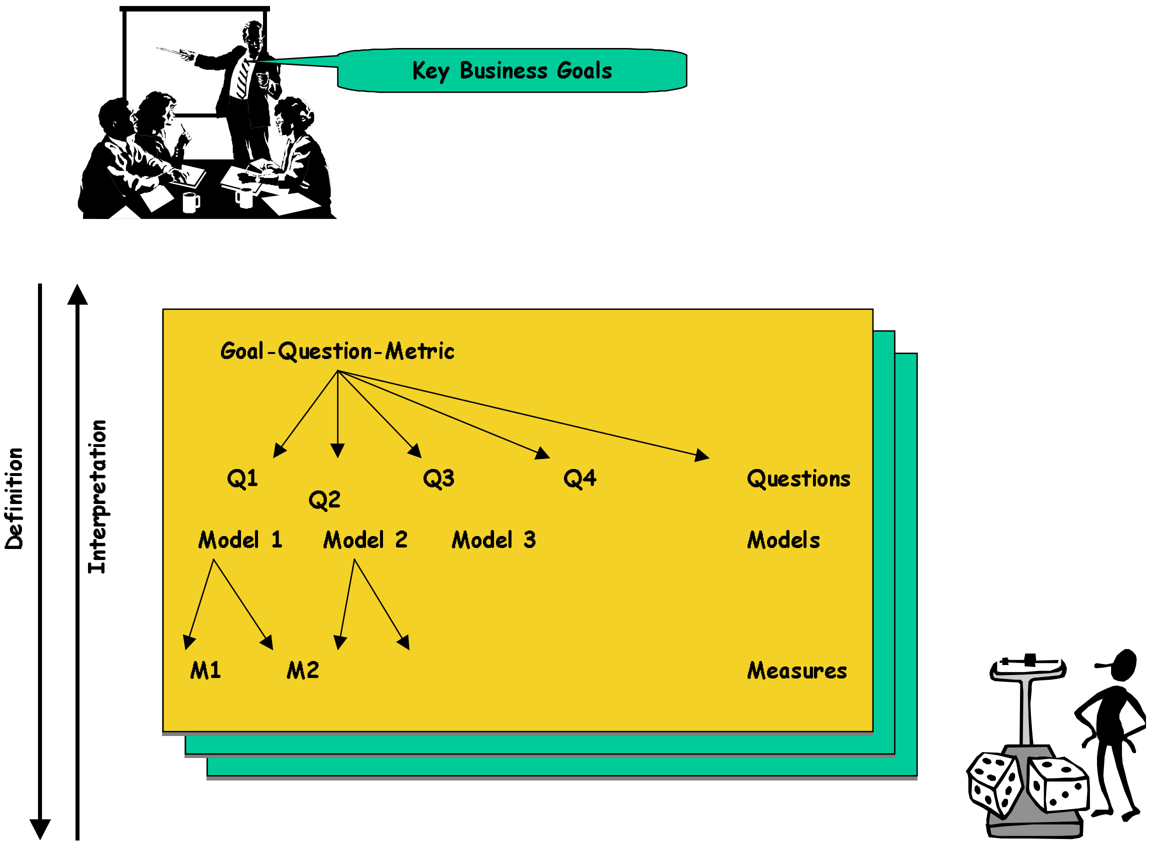


Fig. 4.1 The Basic principle of Goal Question Metric methodology

If ontologies are to be used successfully in industrial settings they should aim to satisfy a number of criteria that reflect basic requirements of the target user communities.
To date a number of researchers/organizations have published such criteria and in this section we will present a set of criteria that approaches in our view what might be a consensus view.

In evaluating ontologies it is useful to organize criteria into categories that will allow the interested party to compare options in a consistent and meaningful manner. For a finer-grained evaluation, the categories themselves can be further grouped according to the envisioned use of the ontology.

4.2 Generic Criteria

Based on the discussion of the previous sections we propose the following categories for evaluation:

- **Modeling Capabilities**: refers to the expressive richness of the underlying language/formalism
- **Supporting tools**: refers to infrastructure issues that make the use of the ontology viable
- **Performance**: a measure of the behavior in terms of computing resources required and kinds of queries available
- **Practicalities**: deployment issues such as scope of use, external connectivity, support

We present the criteria within these categories first in the context of the use of an ontology as a representation medium. We then consider each of these categories in the context of information retrieval.

Modeling Capabilities

- Does the language support the following?
  - ✓ Multiple inheritance
  - ✓ Value constraints
  - ✓ Default values
  - ✓ Conjunction, disjunction, negation
  - ✓ Methods for value calculation
  - ✓ Multi-valued properties
- What primitive data types (e.g. numbers) does the language support
- Does the language support instances as well as classes?
- If the language supports methods and/or constraints how rich is the formalism?
- Does the language support multiple versions of data? Does the language support time stamping of data?

Supporting tools

- What tools exist for maintaining the ontology?
- Are the tools themselves maintained?
- What types of queries can be expressed by the tools?
- Are there any limits in the size of the ontology, names, values etc.?
- What kinds of reasoning mechanisms does the ontology support?

Performance

- What queries can be expressed in the language?
  - ✓ Parents/Children of Concept X
  - ✓ Least common parent of concepts X and Y
  - ✓ Closest common child of concepts X and Y
  - ✓ Links between concepts X and Y
- Is there a language for expressing these queries?
- What are the computing resources requirements of the language?
- Is the tractability proved or bounded?
- Is there reasoning services developed and available?

Practicalities

- Is the ontology being used in any real applications? How large?
- How widespread is its use?
- Is any industry-wide body backing it?
- Links with other ontologies: Are there tools that allow the ontology to be linked with other ontologies?
- What support is available for the ontology formalism (manuals, tutorials, support team)?
- Does a stable release (not to change within the next 6 months) exist?

4.3 References

[Rombach, 1991], Rombach H.D. , **“Practical benefits of goal-oriented measurement”,** in Fenton and Littlewood editors, Software Reliability and Metrics, pages 217-235. Elsevier Applied Science, London.
[Menzies, 1998] T. Menzies, "**Requirements for good measurements**", http://www.cse.unsw.edu.au/~timm/pub/eval/cautions.html
[Julieanne van Zyl, Dan Corbett, 2000 a] **"Population of a Framework for Understanding and Classifying Ontology Applications"**, *ECAI'2OOO*
[Julieanne van Zyl, Dan Corbett, 2000 b] **"A Framework for Comparing the use of a Linguistic Ontology in an Application"** , *ECAI'2000*
[Nick 1999], Nick M., Althoff K., Tauz C. , **“Facilitating the Practical Evaluation of Knowledge-Based systems and Organisational Memories using Goal-Questions-Metric Technique”,**.Proceedings of KAW99
[Briand 1996], Briand L.C., Differding C.M., Rombach H.D. , **“Practical guidelines for measurement-based process improvement”,** Software process, 2(4): pages 253-280

5 Successful Scenarios and Guidelines

5.1 Introduction

Successful scenarios and Guidelines are documented strategies and process employed by top-of-class Applications. One given application can not claimed to be top-of-class in every area - such application does not exist - so in D21 and series, we are capturing the best practices of the top best applications in a given class.

5.2 Corporate Intranet and Knowledge Manag ement

[Rose Dieng 2001] As a conclusion, our survey confirms the multiple research fields relevant for building a corporate KM -which definitively requires a multidisciplinary approach. The choice between the different construction techniques can be based on several questions that an enterprise should answer before building a corporate KM:

1. Needs detection:

- Design a knowledge-management system based on user needs, requirements and usage habits.
- Who are the potential users of the CM and what are the users' profiles?
- What is the intended use of the CM after its construction: is it...
- a way of communication between distant groups?
- a way of communication between an enterprise and privileged partners?
- a way to enhance learning of new enterprise members?
- When will the CM be used: in short-term, in mid-term, or in long-term?

2. Construction:

- Choose a system that is consistent with your organisation's culture and user patterns.
- Work with existing systems. Do not duplicate existing resources.
- What are the knowledge sources available in the firm: paper-based, semi-structured or structured documents, human specialists, databases?
- Can the quality, volume, availability of the knowledge sources be assessed?
- What is the knowledge map of the enterprise departments involved in the knowledge management operation?
- What kind of knowledge must contribute to the construction of the Corporate KM?:
- knowledge already described in documents such as reports or synthesis document on a project?
- elements of experience and professional knowledge not already described in documents?
- Is it necessary to model knowledge of some enterprise members or is an intelligent documentary system sufficient?
- What is the preferred materialisation, according to the computer environment of bothfuture users and developers and according to the financial, human and technical means available for the CM construction and maintenance?
- Organise information simply. Complexity discourages usage.

- Design a system that adds value to information.
- Set a realistic budget to develop and maintain the system. Then adhere to the budget.

3. Diffusion:

- Seek top management support for the system and ensure a high-level champion actively promotes and encourages use of the system.
- What is the preferred scenario of interaction between the future users and the Corporate KM?
- What interface will be the most adapted to the users' activity environment?
- What will be the privileged diffusion means (Internet, Intranet, ...), according to the computer environment of both future users and developers?
- Encourage and train people to use the system.

4. Evaluation:

- What will be the evaluation criteria ?
- When, how and by whom will such an evaluation be carried out?

5. Evolution:

- How will the evaluation results be taken into account?
- When, how and by whom will the Corporate KM be maintained, verified and incremented?
- How will obsolete or inconsistent knowledge be detected and removed (or contextualised)?
- Will the evolution of the Corporate KM be centralised by a department or will it be distributed among several members of the organisation? In the second case who will coordinate the effort?
- Frequently updated information in the system. Designate an individual to manage and maintain the system.

5.3 E-Commerce

5.3.1 E-Business : B2C, B2B

Companies [Best Practices LLC] that fully integrate the Internet into their overall strategy and existing business operations are experiencing dramatic productivity gains and marketplace advantages. Those managers and companies making dramatic progress in e-marketing, e-sales and e-services -- observe it is only a matter of time before the Internet is completely integrated into every facet of their overall strategy and business operations. Then, these e-commerce leaders reflect, today's "e-business" will be tomorrow's "business as usual." . This study is anchored on four critical perspectives that serve as the cornerstones for field work, analysis and organization of all insights. These critical perspectives include the following:

- The Internet has emerged as an essential tool to enable and supercharge productivity.
- The Internet can and should be used to enhance existing successful business models.
- Internet managers must translate e-business tools, techniques and initiatives into real stakeholder value that can be calculated in traditional terms of business economics and productivity.
- Broad-based Internet strategies and initiatives can be analyzed and understood in terms of specific practices that are repeatable and manageable throughout multiple enterprises.

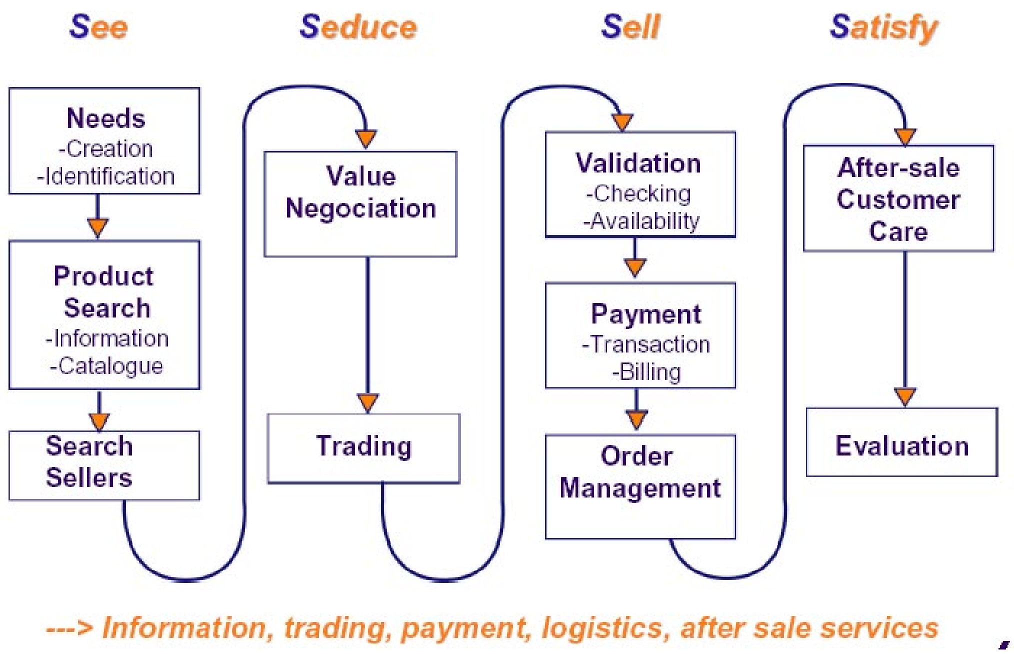
Key Findings

Five Key Findings that have been reported by [Best Practices LLC] follow:

- Evaluate your business model and apply Internet tools to the highest value points.
- Integrate e-tools with traditional customer service excellence principles to drive long-term loyalty, faster responses, greater ease of use, increased functionality and cost reduction.
- Enhance sales productivity by integrating Internet tools with existing sales channels.
- Employ e-marketing to gather information and to segment and target customers with mass-customized communication.
- Measure e-business success based on deployment of people, technology and strategy to accelerate improvement cycles and increase profit margins.

5.3.2 Ontology based B2C

Whilst many private initiatives have used ontologies as part of B2C applications there has, to date, been little effort to formalize best practice within the B2C electronic commerce applications. However, the Ontology Based technology should bring added-value to the classical Value Chain :



5.3.3 Ontology based B2B

[CEN/ISSS] Classification schemes have a long history within electronic commerce. The most influential of these schemes have been developed under the auspices of United Nations' Electronic Data Interchange for Administration, Commerce and Transport (UN/EDIFACT) initiative by the United Nations Centre for Trade Facilitation and Electronic Business (UN/CEFACT).

Other long-standing classification schemes that are relevant to E-Commerce include:

- the Universal Standard Product and Services Classification (UNSPSC) developed by the Electronic Commerce Code Management Association (ECCMA)
- ISO 4217: Codes for the representation of currencies and funds developed by the International Organization for Standardization
- the International Standard Industrial Classification of All Economic Activities (ISIC) developed by the UN Statistics Office.

At present, however, established standardization bodies have not chosen to develop formalized ontologies as basis for classification schemes.

Best practices relating to the use of classification schemes within B2B electronic commerce are currently being studied and standardized under the auspices of UN/CEFACT and OASIS as part of the ebXML initiative and the follow-on ebTWG and UBL work. Details of this on-going work can be found at http://www.ebtwg.org and http://www.oasis-open.org/committees/ubl/.

A wide range of industry specific ontologies are being prepared, among the more advanced of which are the ones specific to the electronic component supply industry (RosettaNet) and the automotive industry (AIAG). Details of the current practices of both these groups can be found at their websites, http://www.rosettanet.org and http://www.aiag.org.

[Feindt et al. 2001] Once an SME has drawn up a strategy for supporting B2B e-commerce interactions, it can determine the factors critical to success in meeting this strategy. Critical success factors are defined as 'the limited number of areas in which results, if they are satisfactory, will ensure successful competitive performance for the organisation.' [Chappell, 2000] Based on previous research and the 43 SMEs of the KITS sample we have identified the following six critical success factors:

- **Vision**: Understanding that an SME's competitiveness in its industry value chain depends on external efficiencies as well as internal ones – that is, efficient interactions with customers and suppliers
- **Culture**: Encouraging open access to information and collaborative processes across internal and external boundaries
- **Trust**: Encouraging value chain partners to feel as high a level of trust in the SME as they would in their own organisations
- **Value**: Ensuring that the B2B e-commerce interaction between value activities (for example, a customer's procurement activity and a supplier's sales and marketing activity) is innovative and adds value over traditional forms of interaction (such as paper, phone or fax)
- **Control**: Establishing the appropriate level of control over the value chain interaction, depending on factors such as business environment, technological capability of the partners, critical nature of the value chain interaction
- **Integration**: Providing internal links between the organization's IT systems and between its IT systems and those of its partners

5.4 Information Retrieval

The last few years have seen a dramatic increase in the generation of information as a result of high throughput technologies in certain leading edge fields (e.g. in the genomics field with the reading of the human genome) but also as a consequence of the ease with which information can be published.

A result of this explosion is that the extraction of knowledge from information, and not the availability of information itself, has become a critical factor for business. From a business perspective therefore, effective information retrieval has become an essential element of the competitive capability of companies within knowledge-intensive industries. Furthermore, the central role that information plays in decision-making creates the need for high quality of the retrieved information.

Most technologies that have been used to date have paid lip service to quality issues partly because other performance metrics (such as response times) were of higher priority. This however is no longer the case and instead quality issues (such as relevance, accuracy, completeness, conciseness) are now important.

Ontology-based approaches promise to increase the quality of responses since they aim to capture within computer systems some part of the semantics of concepts allowing for better information retrieval. However, while the opportunities for value creation exist, the path is anything but straight.

Amongst the opportunities and risks that exist, the main ones are the following:

Opportunities

- Better information retrieval means better decision-making
- Identification of business opportunities (matching buyers with sellers)
- Enhanced scientific discovery capability
- Enhanced problem solving capability

Risks

- The use of proprietary ontologies for all critical applications creates the need for high quality ontology alignment. Failure to achieve this will put the vision of system interoperability at risk.
- Ontologies must be flexible, open ended and capable of dealing with multiple definitions. Failure to achieve this will jeopardize the acceptance of ontology-based systems.

Information retrieval specific requirements

The scenario described in the preceding section poses a set of requirements for ontology-based information retrieval systems. The following is a list of criteria specific to information retrieval:

- What are the precision and recall of the system?
- What kinds of relevance criteria are being used on the query results? In the case of web pages is it number of hits, number of links or some other measure?
- How timely is the information returned?
- How transparent is the system? (i.e. can one drill down to the source?)
- Does the system indicate possible contradictions or multiple versions of the same information?
- Does the system allow for any kind of trend analysis or detection in the query results?

- Does the ontology allow for the representation of context?
- What are the special features of the query language? Does it support the augmenting of the user's query?
- Is there a visualization system available that supports query formation?
- Does the system allow for different kinds of search modes e.g. specific search, browsing, discovery?

Guidelines

The aim of this section is to build on the framework discussed in previous sections and provide a practical checklist of issues to examine when considering the implementation of an ontology-based application.

#	ITEM	ISSUES
1	**Decide on the nature of the application**	Will the ontology be used for ✓ Descriptions? ✓ Integration of heterogeneous info systems? ✓ Better information retrieval? ✓ Support of problem solving? Based on the intended use a different set of evaluation criteria should apply
2	**Identify users and involve them from start**	Because ontologies are about definitions and 'understanding' between communicating parties, where humans are to be involved they constitute a fundamental issue. To improve *transparency* and hence acceptance of the system, its intended users must be consulted from the very beginning.
3	**Look at company information/knowledge 'silos'**	Often groups within large organizations exist in 'silos' where communication between them is non-existent or not very effective. If this is the case, are there duplicate databases/ontologies and how can the ontology application help bridge the gaps between the silos?
4	**Decide on proprietary and non-proprietary concepts (parts of ontology)**	If the intended application is of strategic importance, are there some parts of the ontology that must be kept company confidential? If this is the case, is a partitioning between the proprietary and public parts of the ontology possible? Can the non-proprietary part be found in the public domain and what are the implications in terms of integration with the proprietary part?
5	**Examine your legacy systems**	If the ontology application is not custom built check that integration with legacy systems is possible. This should go beyond the communications layer and consider the logical layers of the legacy and ontology systems.
6	**Decide on Reasoning Processes**	What kinds of use will the ontology be put to? If it to be used for reasoning and problem solving, are there any reasoning engines available or would it be necessary to develop one in-house? What is the state of the art and what the level of maturity of any existing technologies?

7	**Research Existing Ontologies**	Does an existing ontology exist that covers your application needs? This possibility should be considered seriously if the application will need to make use of industry wide information norms and if the information is not considered of strategic value to the company.
8	**Look at Ontology Maintenance tools and procedures**	If a specific ontology environment (that accompanies the ontology) is to be used examine the capabilities of the supporting tools, the assumptions of the underlying language and its expressive power. How easy is it to change concept definitions? Does the environment check for inconsistencies? Can it cope with multiple versions? etc.
9	**If you are to use external ontologies, look at who maintains/backs them**	Is the external ontology you have decided to use approved/backed by any industry wide body? Is there a maintenance procedure? Is there a working group that maintains the ontologies and allows users to voice their needs? Are other industry participants using this ontology?
10	**Consider integration with your supplier/customer applications**	Is there any benefit in integrating the ontology application with that of your customers or suppliers? If so, do they have any equivalent applications in place? Is integration with your ontology possible and at what level?

Table 5.4 : Checklist for the implementation of an ontology application

5.5 Portals and Web communities

Lessons Learnt from the SEAL Representative application [Stajanovic et al. 2001]

The conceptual backbone of the SEAL approach is an ontology. For the AIFB portal, they had to model concepts relevant in this setting. As SEAL has been maturing, a methodology was developed for setting up ontology-based knowledge systems [Staab S. et al. 2001]. The approach (*cf.* Figure Ontology development) is mainly based on [Schreiber B. et al. 1999] and [Gomèz-Pérèz A. 1996] but focuses on the application-driven development of ontologies. Here is described some experiences made during the ontology development for the SEAL portal.

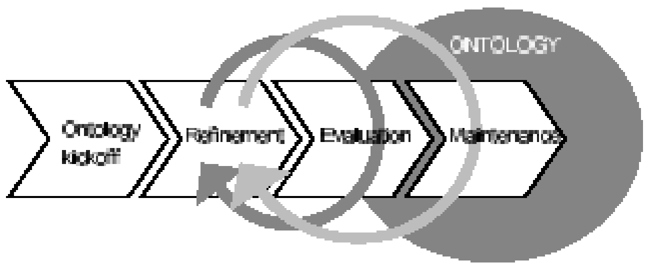


Figure 5.5.1 Ontology Development : global picture

Kickoff phase for ontology development.
Setting up requirements for the SEAL ontology the SEAL team had to deal mainly with modeling the research and teaching topics addressed by different groups of the AIFB institute and personal information about members of the AIFB institute. The SEAL team took themselves as an "input source" and collected a large set of lexical entries for research topics, teaching related topics and personal information, which represent the lexicon component of the ontology. By the sheer nature of these lexical entries, the ontology developers were not able to come up with all relevant lexical entries by themselves. Rather it was necessary to go through several steps with domain experts (*viz.* institute colleagues) in the refinement phase.

Refinement phase.
We started to develop a baseline taxonomy that contained a hierarchy of research topics identified during the kickoff phase. An important result for the team was to recognize that categorization was not based on an isA-taxonomy, but on a much weaker HASSUBTOPIC relationship. *E.g.* "Knowledge Discovery in Data Bases" is a subtopic of "Knowledge Management", which means that it covers some aspects of "Knowledge Management" —but it does not reflect inheritance provided by an isA-taxonomy.

It then took the team three steps to model the currently active ontology. In the first step, lexical entries were collected by all members from the institute. Though the team had already given the possibility to provide a rough categorization, the categories modeled by non-knowledge engineers were not oriented towards a model of the world, but rather towards the way people worked in their daily routine. Thus, their categorization reflected a particular rather than a shared view onto the domain. A lesson learned from this was that people need an idea about the nature of ontologies to make sound modeling suggestions. It was very helpful to show existing prototypes of ontology-based systems to the domain experts.

In the second step, the team worked towards a common understanding of the categorization and the derivation of implicit knowledge, such as "someone who works in logic also works in theoretical computer science" and inverse of relations, *e.g.* "an author has a publication" is inverse to "a publication is written by an author".

In the third step, the team mapped the gathered lexical entries to concepts and relations and organized them at a middle level. Naturally, this level involved the introduction of more generic concepts that people would usually not use when characterizing their work (such as “optimization”), but it also included “politically desired concepts”, because one own’s ontology exhibits one’s view onto the world. Thus, the ontology may become a political issue!

Modeling during early stages of the refinement phase was done with pen and paper, but soon the team took advantage of the ontology environment OntoEdit [see Onto-To-Knowledge] that supports graphical ontology engineering at an epistemological level as well as formalization of the ontology. Like in other ontology engineering projects, the formalization of the ontology is a non-trivial process.

Evaluation phase.

After all the team found that participation by users in the construction of the ontology was very good and met the previously defined requirements, as people were very interested to see their work adequately represented. Some people even took the time to learn about the Ontology editing tool "OntoEdit". However, the practical problem the team had was that our environment does not yet support an ontology management module for cooperative ontology engineering.

We embedded the ontology into our institute portal. It contains around 170 concepts (including 110 research topics) and 75 relations. This version is still running, but we expect maintenance to be a relevant topic soon. Therefore the team are collecting feedback from our users —basically colleagues and students from the institute.

5.6 References in Guidelines and Lessons

Knowledge Management

[Dieng et al., 1999], R. Dieng, O. Corby, A. Giboin, and M. Ribiere. **Methods and tools for corporate knowledge management**. Int.Journal of Human-Computer Studies, 51(3):567–598, 1999.

Information Retrieval

[Qi Li et al 2001], Qi Li, Philip Shilane, Natalya Fridman Noy, M.A. Musen "**Ontology Acquisition from on-line Knowledge Sources**", AMIA Inc. pp. 497 - 501

[L. Crow 2001], L. Crow, N. Shadbolt "**Extracting focused knowledge from the semantic web**" Int. J. Human-Computer Studies 54, pp 155-184

[S. A. McIlraith 2001], S.A. McIlraith, T.C Son, H. Zeng "**Semantic Web Services**". IEEE Intelligent Systems, pp. 46-53

E-business, E-commerce

[E-work and E-commerce 2001] "**Novel solutions and practices for a global networked economy**", e.2001, 17-20 October 2001, Venice, Italy. Edited by Brian Stanford-Smith and Enrica Chiozza, ISBN: 1 58603 205 4.

URL: http://www.ebew.net/programme.htm

[Chappell, 2000], C. Chappell, S. Feindt "**Analysis of E-Commerce prcatice in SMEs**", Communications and Strategies, N°37, 1st trimester 2000

[Feindt et al. 2001], Sylvie Feindt et al., "**Critical Success Factors for eBusiness for SMEs",** in [E-work and E-commerce 2001]

Best Practices, LLC http://www.best-in-class.com

Portals for Web Community

[Staab S; et al. 2001] S. Staab, H.-P. Schnurr, R. Studer, and Y. Sure**. "Knowledge processes and ontologies**". *IEEE Intelligent Systems*, 16(1):26–34, 2001.

[Schreiber G. et al. 1999] G. Schreiber, H. Akkermans, A. Anjewierden, R. de Hoog, N. Shadbolt, W. Van de Velde, and B. Wielinga. *Knowledge Engineering and Management — "***The CommonKADS Methodolog***y"*. The MIT Press, Cambridge, Massachusetts; London, England, 1999.

[Gomez-Pérèz A. 1996] A. Gomez-Perez. "**A framework to verify knowledge sharing technology**". *Expert Systems with Application*, 11(4):519–529, 1996.

6 Tools and Methodology for Ontology-based Applications

6.1 Ontology engineering[7]

Ontologies aim at capturing domain knowledge in a generic way and provide a commonly agreed understanding of a domain, which may be reused and shared across applications and groups. Ontologies typically consist of definitions of concepts, relations and axioms. Until a few years ago the building of ontologies was done in a rather *ad hoc* fashion. Meanwhile there have been some few, but seminal proposals for guiding the ontology development process (*cf.6.1.2*).

In this chapter we describe a methodology for building an ontology-based application by introducing specific guidelines for developing and maintaining the respective ontology. Special emphasis is put on a stepwise construction and evaluation of the ontology.

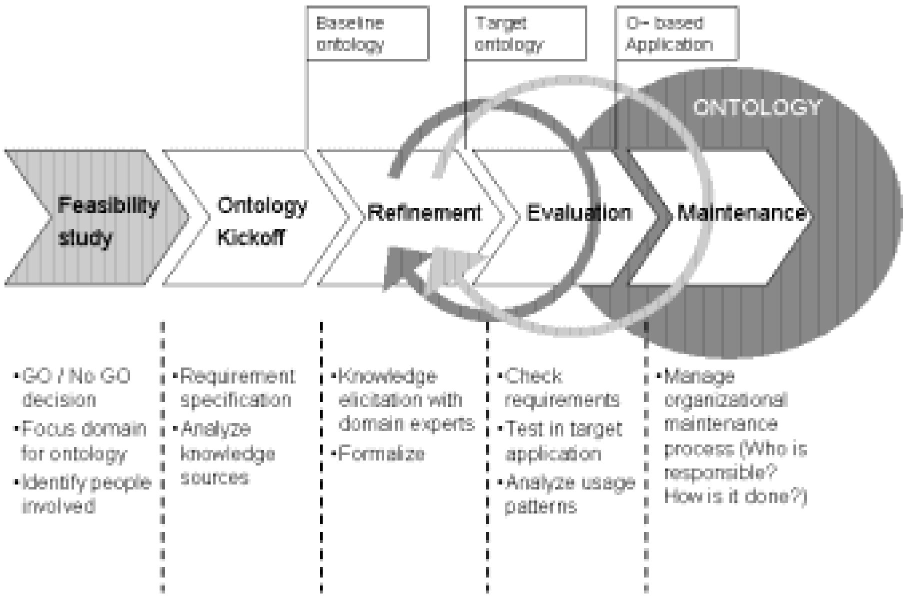


Figure 6.1 Steps of the Ontology Building methodology

6.1.1 Methodology

Kick-off phase. The first step to actually engineer ontologies is to capture requirements in an Ontology Requirements Specification Document ("ORSD") describing what an ontology should support and sketching the planned area of the ontology application. It should guide an ontology engineer to decide about inclusion, exclusion and the hierarchical structure of concepts in the ontology. In this early stage one should look for already developed and potentially reusable ontologies. In summary, it should clearly describe the information shown in Table 6.1.

Through analysis of the available knowledge sources a "baseline ontology" is gathered, *i.e.* a draft version containing few but seminal elements of an ontology. Typically the most important concepts and relations are identified on an informal level. A very important knowledge source (also for the later phases) are domain experts. There exist several possibilities to capture knowledge from domain experts, we focus on the usage of informal competency questionnaires as proposed by [Bernaras A. 1996]. They

[7] The content is extracted from [Sure Y. 2002]

consist of possible queries ("competency questions") to the system, indicating the scope and content of the ontology.

ORSD1. Domain and goal of the ontology
ORSD2. Design guidelines to ensure a consistent development (*e.g.* naming conventions)
ORSD3. Available knowledge sources (*e.g.* domain experts, reusable ontologies, organization charts, business plans, dictionaries, index lists, db-schemas etc.)
ORSD4. Potential users and use cases
ORSD5. Applications supported by the ontology

Table 1 Content of the ORSD

Refinement phase. The goal of the refinement phase is to produce a mature and application-oriented "target ontology" according to the specification given by the kickoff phase. This phase is divided into different sub phases shown in Table 6.2.

R1. A knowledge elicitation process with domain experts based on the initial input from the kickoff phase. This serves as input for further expansion and refinement of the baseline ontology. Typically axioms are identified and modeled in this phase. This is closely linked to the next step – the effects of axioms might depend on the selection of the representation language.
R2. A formalization phase to transfer the ontology into the "target ontology" expressed in formal representation languages like DAML+OIL [DAML 2001]. The representation language is chosen according to the specific requirements of the envisaged application.

Table 6.2 Two sub phases of the refinement phase

This phase is closely linked to the evaluation phase. If the analysis of the ontology in the evaluation phase shows gaps or misconceptions, the ontology engineer takes these results as an input for the refinement phase. It might be necessary to perform several iterative steps.

Evaluation phase. The evaluation phase serves as a proof for the usefulness of developed ontologies and their associated software environment. In a first step, the ontology engineer checks, whether the target ontology fulfils the ontology requirements specification document and whether the ontology supports or "answers" the competency questions analyzed in the kick-off phase of the project. In a second step, the ontology is tested in the target application environment. Feedback from beta users may be a valuable input for further refinement of the ontology.

A valuable input may be as well the usage patterns of the ontology. The prototype system has to track the ways users navigate or search for concepts and relations. With such an "ontology log file analysis" we may trace what areas of the ontology are often "used" and others which were not navigated. As mentioned before, this phase is closely linked to the refinement phase and an ontology engineer may need to perform several cycles until the target ontology reaches the envisaged level— the roll out of the target ontology embedded into the ontology-based application finishes the evaluation phase.

Maintenance phase. In the real world things are changing — and so do the specifications for ontologies. To reflect these changes ontologies have to be maintained frequently like other parts of software, too. We stress that the maintenance of ontologies is primarily an organizational process. There must be strict rules for the update-insert-delete processes within ontologies. Most important is to clarify *who* is responsible for maintenance and *how* it is performed. *E.g.* is a single person or a consortium responsible for the maintenance process? In which time intervals is the ontology maintained? We recommend that the ontology engineer gathers changes to the ontology and initiates the switch-over to a new version of the ontology after thoroughly testing possible effects to the application, *viz.* performing additional cyclic refinement and evaluation phases. Similar to the refinement phase, feedback from users may be a valuable input for identifying the changes needed. Maintenance should accompany ontologies as long as they are on duty.

6.1.2 Related Work

We here give an overview of existing methodologies for ontology development and show which of their ideas are adopted and expanded in our methodology.

Skeletal Methodology. This methodology is based on the experience of building the Enterprise Ontology (*cf*. [Ushold et al. 1995]), which includes a set of ontologies for enterprise modeling. The guidelines for developing ontologies start with identifying the purpose of an ontology and then concentrate on the building of ontologies which is broken down into the steps ontology capture, coding, evaluation and documentation. A disadvantage of this methodology is that it does not precisely describe the techniques for performing the different activities. For example, it remains unclear, how the key concepts and relationships should be acquired. Only a very vague guideline is given.

We catch up the idea of competency questions and expand their usage. We not only propose to use them for evaluation of the system, but also for finding relevant lexical entries like concepts, relations etc..

KACTUS. The approach of [Bernaras A. 1996] was developed within the Esprit KACTUS project. One of the objectives of this project was to investigate the feasibility of knowledge reuse in complex technical systems and the role of ontologies to support it. The methodology recommends an application driven development of ontologies. So, every time an application is assembled, the ontology that represents the knowledge required for the application is built. Three steps have to be taken every time an ontology-based application is assembled:

1. Specification of the application. Provide an application context and a view of the components that the application tries to model.
2. Preliminary design. Based on relevant top-level ontological categories create a first draft where the list of terms and application specific tasks developed during the previous phase is used as input for obtaining several views of the global model in accordance with the top-level ontological categories determined. Search for existing ontologies which may be refined and extended for use in the new application.
3. Ontology refinement and structuring. Structure and refine the model in order to arrive at a definitive design.

The methodology offers very little detail and does not recommend particular techniques to support the development steps. Also, documentation, evaluation and maintenance processes are missing [7]. In general we agree with the general idea of application driven ontology development and in particular with refinement and structuring, which is reflected by our proposal of the ontology development process.

Methontology. The Methontology framework from [Gomez A. 1998] includes:

1. The identification of the ontology development process, which refers to which tasks (planning, control, specification, knowledge acquisition, conceptualisation, integration, implementation, evaluation, documentation, configuration management) one should carry out, when building ontologies.
2. The identification of stages through which an ontology passes during its lifetime.
3. The steps to be taken to perform each activity, supporting techniques and evaluation steps.
4. Setting up an ontology requirements specification document (ORSD) to capture requirements for an ontology similar to a software specification.

The methodology offers detailed support in development-oriented activities except formalization and maintenance and describes project management activities. We adopted the general idea of an ontology requirements specification document (ORSD), but modified and extended the presented version by our own needs.

6.2 Ontology building from text from text[8]

One promising approach for establishing an ontology and acquire knowledge is to incorporate results from disciplines like linguistics. Researchers in terminology for example are interested in organizing domains from a conceptual point of view from the analysis of terms used to name concepts in texts. On the other hand, an ontology is based on the definition of a structured and formalized set of concepts, and a great part of it comes from text analysis, such as transcript of interviews, and technical documentation. In such cases, the theory of a domain can only be found by reaching concepts from terms.

For several years, some researchers in terminology have identified a parallel between terminology as a practical discipline and artificial intelligence, in particular knowledge engineering. From a knowledge engineering point of view, we notice two trends. One trend is to propose to elicit knowledge by using automatic processing tools, widely used in linguistics. Another one is to establish a synergy between research works in artificial intelligence and in linguistics, by means of terminology. An overview of these developments is given below.

Natural language processing tools may help to support modeling from texts in two ways. First, they can help to find the terms of a domain [Bou94], [BGG96] [OFR96]. Existing terminologies or thesauri may be reused and increased or new ones may be created. Second, they can help to structure a terminological base by identifying relations between concepts [Jou95] [JME95] [Gar97].

Three steps are necessary to find the terms of a domain. At the beginning, nominal groups are isolated from a corpus considered as being representative of the studied domain. Then, those that can't be chosen as terms because of morphological or semantic characteristics are eliminated. Finally, the nominal sequences that will be retained as terms are chosen. Usually, this last step requires a human expertise.

Identifying relations between concepts is composed of three steps too. The first one identifies the co-occurrences of terms. Two terms are co-occurrent if they both appear in a given text window which may be defined in several ways: a number of words, a documentary segmentation (entire document, section), a syntactic cutting of sentences, ... The second step computes a similarity between terms with respect to contexts they share. Then, the third step can determine the terms that are semantically related. In most cases, identified relations are the following: semantic proximity, meronimy, causal or more specific relations.

Some researchers have focussed on trying to benefit from approaches from both linguistics and knowledge engineering. They have studied mutual contributions, and their work has led them to elaborate the concept of Terminological Knowledge Base (TKB). This concept was first defined by Ingrid Meyer [SMe91] [MSB+92].

Building a TKB is seen as an intermediate model that helps toward the construction of a formal ontology. A TKB is a computer structure that contains conceptual data, represented in a network of domain concepts, but also linguistic data on the terms used to name the concepts. Thus a TKB contains three levels of entities: term, concept and text. It is structured by using three kinds of links. Relations between term and concept allow synonymy and paronimy to be considered. Relations between concepts compose the network of domain concepts. Relations between term and/or concept and text allow normalization choices to be justified or knowledge base to be documented. A TKB is interesting to build a KBS, especially because it gathers some linguistic information on terms used to name concepts on. This can enhance communication between experts, knowledge engineers and end-users, or be a great help for the knowledge engineer to choose the names of the concepts in the system. Nevertheless, if most researchers agree with its structure, problems still remain today about genericity and also about the construction and the exploitation of the corpus, which is very important in the construction of the TKB because it is the reference from which modeling choices will be justified. Current research continues in these directions.

---

[8] Contribution from Univ. d'Orsay, Paris Sud, LRI (Chantal Reynaud) as in FIPA Part 12 Ontology services (Annex B Guidelines to define a new Ontology under editorship of A. Léger FIPA 1998]

6.3 References

[Ushold M. 1995] M. Uschold and M. King. "**Towards a Methodology for Building Ontologies**". In *Workshop on Basic Ontological Issues in Knowledge Sharing, held in conjunction with IJCAI-95*, Montreal, Canada, 1995.

[Gomez 1996] A. Gomez-Perez. "**A Framework to Verify Knowledge Sharing Technology**". *Expert Systems with Application*, 11(4):519– 529, 1996.

[Ushold M. 1996] M. Uschold & M. Grüninger, "**Ontologies: Principles, methods and applications**" Knowledge Sharing and Review, 11(2)

[Bernaras A. et al. 1996] A. Bernaras, I. Laresgoiti, and J. Corera**. "Building and Reusing Ontologies for Electrical Network Applications**". In *Proceedings of the European Conference on Artificial Intelligence ECAI-96*, 1996.

[Gomez-Perez A. 1998] Assuncion Gomez-Pérez, "**Methontology" « Knowledge Sharing and Reuse »,** Laboratorio de Intelligencia Artificial, Facultad de informatica, Universidad Politécnica de Madrid.

[D. O'Leary 1998] D. O'Leary. "**Using AI in knowledge management: Knowledge bases and ontologies**". *IEEE Intelligent Systems*, 13(3):34– 39, May/June 1998.

[Lopez F 1999] F. L´opez. "**Overview of methodologies for building ontologies**". In *Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving Methods: Lessons Learned and Future Trends*. CEUR Publications, 1999.

[Sure Y. 1999] York Sure, Rudi Studer, "**On-To-Knowledge Methodology - Employed and Evaluated Version**", On-To-Knowledge Project IST-1999-10132, pp. 23-30

[Staab S 2001] S. Staab, H.-P. Schnurr, R. Studer, and Y. Sure. "**Knowledge processes and ontologies**". *IEEE Intelligent Systems*, 16(1):26– 34, 2001.

[Sure Y. 2001 Y. Sure and R. Studer. "**On-To-Knowledge Methodology — Evaluated and Employed Version**". On-To-Knowledge deliverable D-16, Institute AIFB, University of Karlsruhe, 2001.

[Sure Y. 2002] York Sure, "**A Tool-supported Methodology for Ontology-based Knowledge Management**", submitted to ISMIS 2002, Methodologies for Intelligent Systems

[DAML 2001] **DAML+OIL**. http://www.daml.org/2001/03/daml+oil-index

[BCo95] Bourigault D., Condamines A., "**Réflexions autour du concept de base de connaissances Terminologiques**", Dans les actes des journées nationales du PRC-IA, Nancy, 1995.

[Bou94] Bourigault D., "**LEXTER, un logiciel d'extraction de terminologie. Application à l'acquisition des connaissances à partir de textes**", Thèse de l'Ecole des Hautes Etudes en Sciences Sociales (Paris).

[BGG96] Bourigault D., Gonzalez-Mullier I., Gros C**., "LEXTER, a natural Language Processing Tool for Terminology Extraction**", actes de EURALEX'96 (Göteborg).

[Gar97] GARCIA D., **"COATIS, an NLP System to Locate Expressions of Ations Connected by Causality Links**", in Proc. 10th European Workshop, EKAW'97, San Feliu de Guixols, Catalonia, Spain, October 97, LNAI 1319, pp. 347-352, 1997.

[Jou95] Jouis Ch., "**SEEK, un logiciel d'acquisition des connaissances utilisant un savoir linguistique sans employer de connaissances sur le monde externe**", Actes des 6èmes Journées Acquisition et Validation (JAVA'95), Grenoble, pp. 159-172, 1995.

[JME95] Jouis Ch., Mustafa-Elhadi W., "**Conceptual Modeling of database Schema using linguitic knowledge. Application to terminological Knowledge bases**", First Workshop on Application of Natural language to Databases (NLDB'95), Versailles, Juin 95, pp. 103-118, 1995.

[MSB+92] Meyer I., Skuce D., Bowker L., Eck K., "**Toward a new generation of terminological resources: an experiment in building a terminological knowledge base**. In Proc. 14th International Conference on Computational Linguistics. Nantes. pp. 956-960, 1992.

[OFR96] Oueslati R., Frath P., Rousselot F., "**Term identification and Knowledge Extraction**", International Conference on Applied Natural Language and Artificial Intelligence. Montreal. Juin 96

7 Conclusions

This deliverable is the **first of a series of 5 documents** to be delivered on a **6 monthly basis** and whose target is to report **Best Practices and to give Guidelines for the application of Knowledge IT** to practitioners in the field of - in large - E-Commerce and Knowledge Management.

In this first report, we have tried **to sketch the subject** of "Evaluation and emerging Best Practices" in a so lively, young and promising field of Ontology-based Solutions.

The next releases will not try to extend the length of the document but as Semantic Web promises, to be more and more accurate in setting the Evaluation Framework, selecting the Key or Killer Applications and eventually provide the known Best Practices to **assist anyone in the industry or the commercial sectors to self appropriate rapidly these emerging technologies to answer practical needs.**

8 References

Evaluations

[Rombach, 1991], Rombach H.D. , **“Practical benefits of goal-oriented measurement”,** in Fenton and Littlewood editors, Software Reliability and Metrics, pages 217-235. Elsevier Applied Science, London.

[Grüninger et al.] Michael Grüninger, Mark S. Fox, **"Methodology for the Design and Evaluation of Ontologies",** Proceedings of IJCAI'95, Workshop on Basic Ontological Issues in Knowledge Sharing, April 13, 1995

[Cohen, 1995] **Empirical Methods for Artificial Intelligence**, MIT Press

[Briand 1996], Briand L.C., Differding C.M., Rombach H.D. , **“Practical guidelines for measurement-based process improvement”,** Software process, 2(4): pages 253-280

[Menzies, 1997] T. Menzies, "**Evaluation Issues for Problem Solving Methods**" , Banff KA workshop 1998

[Menzies, 1998] T. Menzies, "**Evaluation Issues with Critical Success Metrics**", Proceedings of KAW’98

[Menzies, 1998] T. Menzies, "**Requirements for good measurements**", http://www.cse.unsw.edu.au/~timm/pub/eval/cautions.html

[Nick 1999], Nick M., Althoff K., Tauz C. , **“Facilitating the Practical Evaluation of Knowledge-Based systems and Organisational Memories using Goal-Questions-Metric Technique”,**.Proceedings of KAW99

[Jasper et al. 1999] R. Jasper and M. Uschold, " **A framework for Understanding and Classifying Ontology Applications** ", Proceedings of KAW’99

[Benjamin et al. 1999], *Asuncion Gomez Perez, V. Richard Benjamins*, **"Overview of Knowledge, Sharing and Reuse Components: Ontologies and Problem-Solving Methods**". In Proceedings of the IJCAI-99 workshop on Ontologies and Problem-Solving Methods (KRR5) Stockholm, Sweden, August 2, 1999. pp. 1.1--1.15

[Sure et al. 1999] York Sure, Rudi Studer, "**On-To-Knowledge Methodology - Employed and Evaluated Version**", On-To-Knowledge Project IST-1999-10132

[Rausher 1999] J. Rauscher, S. Marc, "**A la conquête de la Silicon Valley**" , Editions d’Organisation, Paris, ISBN: 2-7081-2282-7

[Julieanne van Zyl, Dan Corbett, 2000 a] **"Population of a Framework for Understanding and Classifying Ontology Applications"**, *ECAI’2OOO*

[Julieanne van Zyl, Dan Corbett, 2000 b] **"A Framework for Comparing the use of a Linguistic Ontology in an Application"** , *ECAI’2000*

[Nick 1999], Nick M., Althoff K., Tauz C. , **“Facilitating the Practical Evaluation of Knowledge-Based systems and Organisational Memories using Goal-Questions-Metric Technique”,**.Proceedings of KAW99

Knowledge Management

[Staab et al., 2001], Steffen Staab, Rudi Studer, Hans-Peter Schnurr, York Sure, **“Knowledge Management : Knowledge Processes and Ontologies”**; IEEE Intelligent Systems Journal, Jan. Feb. 2001.

[Dieng et al., 1999], R. Dieng, O. Corby, A. Giboin, and M. Ribiere. **Methods and tools for corporate knowledge management**. Int.Journal of Human-Computer Studies, 51(3):567–598, 1999.

[Kuehn et al., 1997], O. Kuehn and A. Abecker. **Corporate memories for knowledge memories in industrial practice: Prospects and challenges**. *Journal of Universal Computer Science*, 3(8), August 1997.

[Schreiber et al., 1999], G. Schreiber, H. Akkermans, A. Anjewierden, R. de Hoog, N. Shadbolt, W. Van de Velde, and B. Wielinga. **Knowledge Engineering and Management — The CommonKADS Methodology**. The MIT Press, Cambridge, Massachusetts; London, England, 1999.

[Dyson S. IBM, 2001], S. Dyson, IBM UK, « **B2E Mobile e-Business : Driver, Passenger or Spectator ?** », in [E-work and E-commerce 2001].

[D. fensel, 2001] D. Fensel: "**Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce**". Springer-Verlag, Berlin, 2001.

[D. Fensel et al. 2002] D. Fensel, J. Hendler, H. Lieberman, and W. Wahlster (eds.): "**Semantic Web Technology**", MIT Press, Boston, to appear.

[D. Fensel, 1998], D. Fensel, S. Decker, M. Erdmann, and R. Studer: "**Ontobroker in a Nutshell**" (short paper). In C. Nikolaou et al. (eds.), Research and Advanced Technology for Digital Libraries, *Lecture Notes in Computer Science*, LNCS 1513, Springer-Verlag Berlin, 1998.

[O'Leary, et al. 2001] Dan E. O'Leary and Rudi Studer (eds.): "**Knowledge Management**", IEEE Intelligent Systems 16 (1), January/February 2001 (special issue).

[Staab S. et al. 2001] Steffen Staab and Alexander Maedche, "**Knowledge Portals — Ontologies at Work**". *AI Magazine* 21(2), Summer 2001.

Steffen Staab, Rudi Studer, Hans-Peter Schnurr, York Sure **"Knowledge Management : Knowledge Processes and Ontologies"**;, IEEE Intelligent Systems Journal, Jan. Feb. 2001.

IST Knowledge Management portal , http://www.knowledgeboard.com/

Ontology Building and Checking

[Ushold M. 1995] M. Uschold and M. King. "**Towards a Methodology for Building Ontologies**". In *Workshop on Basic Ontological Issues in Knowledge Sharing, held in conjunction with IJCAI-95*, Montreal, Canada, 1995.

[Gomez 1996] A. Gomez-Perez. "**A Framework to Verify Knowledge Sharing Technology**". *Expert Systems with Application*, 11(4):519– 529, 1996.

[Ushold M. 1996] M. Uschold & M. Grüninger, "**Ontologies: Principles, methods and applications**" Knowledge Sharing and Review, 11(2)

[Bernaras A. et al. 1996] A. Bernaras, I. Laresgoiti, and J. Corera**. "Building and Reusing Ontologies for Electrical Network Applications**". In *Proceedings of the European Conference on Artificial Intelligence ECAI-96*, 1996.

[Borst P. et al. 1997] Borst P. and Akkermans H. **« Engineering ontologies », Special issue : Using explicit ontologies in knowledge-based system development**, HCS, Vol. 46, Number 2/3, February/March 1997, pp. 365-406.

[Gomez-Perez A. 1998] Assuncion Gomez-Pérez, "**Methontology" « Knowledge Sharing and Reuse »,** Laboratorio de Intelligencia Artificial, Facultad de informatica, Universidad Politécnica de Madrid.

[D. O'Leary 1998] D. O'Leary. "**Using AI in knowledge management: Knowledge bases and ontologies**". *IEEE Intelligent Systems*, 13(3):34– 39, May/June 1998.

[Lopez F 1999] F. L´opez. "**Overview of methodologies for building ontologies**". In *Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving Methods: Lessons Learned and Future Trends*. CEUR Publications, 1999.

[Sure Y. 1999] York Sure, Rudi Studer, "**On-To-Knowledge Methodology - Employed and Evaluated Version**", On-To-Knowledge Project IST-1999-10132, pp. 23-30

[Staab S 2001] S. Staab, H.-P. Schnurr, R. Studer, and Y. Sure. "**Knowledge processes and ontologies**". *IEEE Intelligent Systems*, 16(1):26– 34, 2001.

[Sure Y. 2001 Y. Sure and R. Studer. "**On-To-Knowledge Methodology — Evaluated and Employed Version**". On-To-Knowledge deliverable D-16, Institute AIFB, University of Karlsruhe, 2001.

[Sure Y. 2002] York Sure, "**A Tool-supported Methodology for Ontology-based Knowledge Management**", submitted to ISMIS 2002, Methodologies for Intelligent Systems

[DAML 2001] **DAML+OIL**. http://www.daml.org/2001/03/daml+oil-index

Natural Language based Knowledge acquisition references

[BCo95] Bourigault D., Condamines A., "**Réflexions autour du concept de base de connaissances Terminologiques**", Dans les actes des journées nationales du PRC-IA, Nancy, 1995.

[Bou94] Bourigault D., "**LEXTER, un logiciel d'extraction de terminologie. Application à l'acquisition des connaissances à partir de textes**", Thèse de l'Ecole des Hautes Etudes en Sciences Sociales (Paris).

[BGG96] Bourigault D., Gonzalez-Mullier I., Gros C**., "LEXTER, a natural Language Processing Tool for Terminology Extraction**", actes de EURALEX'96 (Göteborg).

[Gar97] GARCIA D., **"COATIS, an NLP System to Locate Expressions of Ations Connected by Causality Links**", in Proc. 10th European Workshop, EKAW'97, San Feliu de Guixols, Catalonia, Spain, October 97, LNAI 1319, pp. 347-352, 1997.

[Jou95] Jouis Ch., "**SEEK, un logiciel d'acquisition des connaissances utilisant un savoir linguistique sans employer de connaissances sur le monde externe**", Actes des 6èmes Journées Acquisition et Validation (JAVA'95), Grenoble, pp. 159-172, 1995.

[JME95] Jouis Ch., Mustafa-Elhadi W., "**Conceptual Modeling of database Schema using linguitic knowledge. Application to terminological Knowledge bases**", First Workshop on Application of Natural language to Databases (NLDB'95), Versailles, Juin 95, pp. 103-118, 1995.

[MSB+92] Meyer I., Skuce D., Bowker L., Eck K., "**Toward a new generation of terminological resources: an experiment in building a terminological knowledge base**. In Proc. 14th International Conference on Computational Linguistics. Nantes. pp. 956-960, 1992.

[OFR96] Oueslati R., Frath P., Rousselot F., "**Term identification and Knowledge Extraction**", International Conference on Applied Natural Language and Artificial Intelligence. Montreal. Juin 96

[Skuce D. et al 1998] Skuce D., Meyer I., "**Terminology and knowledge acquisition: exploring a symbiotic relationship**". In Proc. 6th Knowledge Acquisition for Knowledge-Based System Workshop, Banff, pp. 29/1-29/21.

[Houssem A. 1998] Houssem Assadi, "**Construction of a regional ontology from text and its use within a documentary system**", FOIS'98, pp. 236-249, Trento, June 98.

E-Commerce and E-Business

[E-work and E-commerce 2001] "**Novel solutions and practices for a global networked economy**", e.2001, 17-20 October 2001, Venice, Italy. Edited by Brian Stanford-Smith and Enrica Chiozza, ISBN: 1 58603 205 4.URL: http://www.ebew.net/programme.htm

[eWork 2001], "Status Report on New Ways to Work in the Knowledge Economy", IST document URL: http://www.eto.org.uk

[MKBEEM 2000], http://www.mkbeem.com/

[MULECO 200], http://www.cenorm.be/isss/Workshop/ec/MULECO/Documents/Muleco_Documents.htm

[SmartEC, 2000], http://www.telecom.ntua.gr/smartec/

[Chappell, 2000], C. Chappell, S. Feindt "**Analysis of E-Commerce prcatice in SMEs**", Communications and Strategies, N°37, 1st trimester 2000

[Feindt et al. 2001], Sylvie Feindt et al., "**Critical Success Factors for eBusiness for SMEs",** in [E-work and E-commerce 2001]

Information retrieval

[Guarino et al., 1999], Nicola Guarino, Claudio Masolo, Guido Vetere**, "OntoSeek: Content-Based Access to the Web**", IEEE Intelligent Systems, 1999.

[Stuart Aitken 2000] Stuart Aitken, Sandy Reid, **"Evaluation of an Ontology-Based Information retrieval Tool"**, ECAI'00, Applications of Ontologies and Problem-Solving Methods.

[Fensel et al., 2000], Dieter Fensel, Stefan Decker, Michael Erdmann, Rudi Studer, "**Ontobroker: Ontology based Access to Distributed and Semi-Structured Information"**
[Qi Li et al 2001], Qi Li, Philip Shilane, Natalya Fridman Noy, M.A. Musen "**Ontology Acquisition from on-line Knowledge Sources**", AMIA Inc. pp. 497 - 501
[L. Crow 2001], L. Crow, N. Shadbolt "**Extracting focused knowledge from the semantic web**" Int. J. Human-Computer Studies 54, pp 155-184
[S. A. McIlraith 2001], S.A. McIlraith, T.C Son, H. Zeng "**Semantic Web Services**". IEEE Intelligent Systems, pp. 46-53

SARI : http://www.ercim.org/publication/Ercim_News/enw35/taveter.html
MCM: http://ai.bpa.arizona.edu/go/intranet/papers/Customizable-00.htm

Portals for Web community

[Stojanovic et al., 2001], N. Stojanovic, A. Maedche, S. Staab, R. Studer, and Y. Sure**. "SEAL— A Framework for Developing SEmantic PortALs"**. In *K-Cap 2001 - First International Conference on Knowledge Capture, Oct. 21-23, 2001*, Victoria, B.C., Canada, 2001. to appear.
URL: http://ontobroker.semanticweb.org/ontos/aifb.html
[Saglio et al. 2002], J. M. Saglio, Tuang Anh Ta, "A Framework for Dynamic Exploration in Semantic Portals", KR2002 to appear
C-Web : Community Webs
URL: http://cweb.inria.fr/
[Staab S. et al. 2000] S. Staab, J. Angele, S. Decker, A. Hotho, A. Maedche, H-P. Schnurr, R. Studer, and Y. Sure. "**AI for the web — ontology-based community web portals**". In *AAAI 2000/IAAI 2000 - Proceedings of the 17th National Conference on Artificial Intelligence and 12th Innovative Applications of Artificial Intelligence Conference*. AAAI Press/MIT Press, 2000.
[Ontobroker, 1999] S. Decker, M. Erdmann, D. Fensel, and R. Studer. "**Ontobroker: Ontology Based Access to Distributed and Semi-Structured Information**". In R. Meersman et al., editors, *Database Semantics: Semantic Issues in Multimedia Systems*, pages 351–369. Kluwer Academic Publisher, 1999.
[SHOE 2000] J. Heflin and J. Hendler. "**Searching the web with SHOE**". In *Artificial Intelligence for Web Search. Papers from the AAAI Workshop. WS-00-01*, pages 35–40. AAAI Press, 2000.
[Staab S; et al. 2001] S. Staab, H.-P. Schnurr, R. Studer, and Y. Sure**. "Knowledge processes and ontologies**". *IEEE Intelligent Systems*, 16(1):26–34, 2001.
[Schreiber G. et al. 1999] G. Schreiber, H. Akkermans, A. Anjewierden, R. de Hoog, N. Shadbolt, W. Van de Velde, and B. Wielinga. *Knowledge Engineering and Management — "***The CommonKADS Methodolog***y"*. The MIT Press, Cambridge, Massachusetts; London, England, 1999.
[Gomez-Pérèz A. 1996] A. Gomez-Perez. "**A framework to verify knowledge sharing technology**". *Expert Systems with Application*, 11(4):519–529, 1996.

General Web Sites

KA Workshops and Archives, http://ksi.cpsc.ucalgary.ca/KAW/
KAW'98 Proceedings, http:// ksi.cpsc.ucalgary.ca/KAW/KAW98/KAW98Proc.html
KAW'99 Proceedings, http://sern.ucalgary.ca/KSI/KAW/KAW99/papers.html
IETF Benchmarking methodology (bmwg); http://www.ietf.org/html.charters/bmwg-charter.html
Best Practices LLC, http://www.best-in-class.com/
ISMBC - Information Systems Management Benchmarking Consortium; http://www.ismbc.org/
IST Knowledge Management portal, http://www.knowledgeboard.com/
SWWS Semantic Web Working Symposium, International Semantic Web Working Symposium, Stanford University, CA USAAugust 1, 2001, http://www.semanticweb.org/SWWS/,