Ontology Domain Modeling Support for Multi-lingual services in E-Commerce : MKBEEM

**Alain Léger , Géraldine Michel , Peter Barrett , Sylvain Gitton, Asuncion Goméz-Pérez , Aarno Lehtola ,
Kristiina Mokkila, Santiago Rodrigez , Jean Sallantin, Theodora Varvarigou, Jérôme Vinesse**

Abstract. One of the main objectives of a truly user-friendly Information Society is to focus on advanced human language technologies enabling cost-effective interchange across language and culture and more natural interfaces to digital services. The recently launched IST-1999-10589 project MKBEEM (Multilingual Knowledge Based European Electronic Marketplace, 1st Feb. 2000 - 1st Aug. 2002) is rightly in that direction and the work will address basically, written language technologies and its use in the key sector of global business and electronic commerce. In particular MKBEEM will focus on adding multilinguality to all stages of the information cycle, including multilingual content generation and maintenance, automated translation and interpretation and enhancing the natural interactivity and usability of the service with unconstrained language input. On the Knowledge engineering side, the MKBEEM Ontologies will provide a consensual representation of the electronic commerce field in three typical Domains (Tourism, Mail order, Retailers) allowing the exchanges independently of the language of the end user, the service, or the content provider. Ontologies will be used for classifying and indexing catalogues, for filtering user's query, for facilitating multilingual man-machine dialogues between user and software agent, and for inferring information that is relevant to the user's request. This paper concentrates on ontology issues, while the used human language processing approaches will be presented closely in our later papers.

1 INTRODUCTION

A key issue for the "user-friendly information society" across Europe is *accessibility* of information - whether by the individual or by business. The Internet, and the raft of multimedia and e-commerce services it has spawned - has provided a physical solution to accessibility which is within the reach of virtually all citizens and businesses, albeit as yet in a somewhat non-integrated nature. However, Europe is made up from (and will continue to be made up from) nationalities with diverse cultures and languages. The information society must develop to support and nurture these rich diversities, rather than act as a constraint to their evolution

through the new media.

At the same time, the growth of a wide range of e-commerce services, both to individuals and between businesses, and the demands of users of these services, are contributing to the increasing international reality of information access and trading of products and services ; suppliers want to manage their product catalogues in their own language, yet want to reach many different nationalities of customer without deploying enormous translation resources ; information providers likewise want to reach their customers irrespective of their language ; many end-users would prefer to use their natural spoken language to communicate with their terminal or computer system ; and service operators, faced with the task of maintaining e-commerce catalogues from multiple suppliers in multiple languages, wonder how they can maintain a reasonably consistent classification space in such a multilingual environment.

So, the global MKBEEM aim is to extend drastically current electronic commerce platforms recently deployed to a broader spectrum of products and services over European countries, to reach a truly pan European and culturally open electronic commerce market. The scientific and technical aims of MKBEEM is to create an *intelligent knowledge based multilingual* mediation service in view of a first proof of concept in the field of European electronic commerce. The **features** of the intelligent knowledge-based multilingual Mkbeem service will allow basically:

- Natural language interface for both the end-user and the Content Provider / Service Provider (CP/SP). Using that interface, end-users will be able to formulate their requests in their own natural language(s), and CP/SP users will be able to formulate their offers in their own natural language(s) and practice(s) ;
- Automatic multilingual cataloguing for the Products and Service Providers ;
- On-line e-commerce contractual negotiation mechanisms in one's language that guarantee safety and freedom.

2 FUNCTIONAL ARCHITECTURE

The global functional architecture of the MKBEEM mediation service appears in figure 1. Let us consider a scenario trained on the functional architecture, making reference to the usage of the different components presented in the MKBEEM architecture.

A Finnish CP / CS user wants to register in MKBEEM new offers. He accesses the **CP Agent** for helping him to enter those new offers in the **Ontological CP catalogue** which describes his own offer. Considering the present partnership this might be

---

related to either a B-to-B catalogue of products, a Mail order consumer products or a railway time table, booking and new special offers. The MKBEEM **Rational Agent** will automatically aggregate / classify this new information in the Ontological CP catalogue by making use of the corresponding **Ontology Domain** server.

An English end user is searching for specific Nordic equipment for country skiing. He then asks his User Agent in English the items he would like to buy. **User Agent** recognises the language in which the user has expressed his need, asks the corresponding HLP Server to map the query to an ontological representation and transmits the ontological query to the **Rational Agent**.

The **Rational Agent** will use this information and the **Ontological CP catalogues** to identify and consult the suitable **CP/SP Agents**, which will in turn provide a list of product items matching the user's description. The Rational Agent will then filter and process this list, with the aid of the **Ontological Contract Management** server, providing information on product availability, applicable taxes etc. Eventually all the details of the **Transactions** are passed to the **E-Commerce Platform** which will incorporate security measures and banking services (not part of this project). Finally, the **Ontological Contract Management** server will incorporate additional conditions if they are required. All these services will be accessible from either of the different interfaces already mentioned.
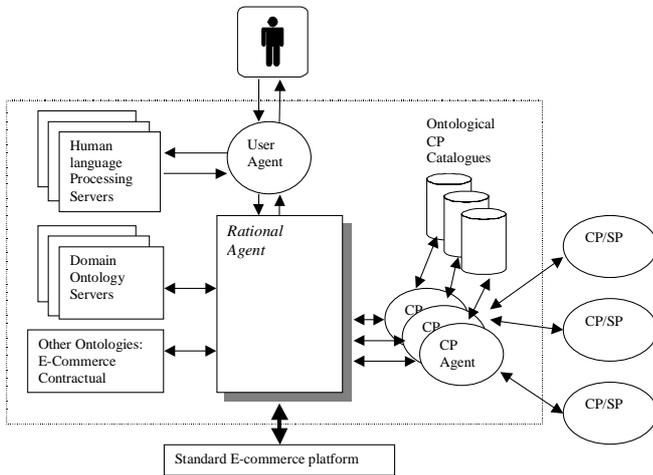

Figure1. Global functional architecture

2.1 The roles of Ontologies

The development of the Internet offers companies opportunities in a broader market place. The main barrier to this development is not so much related to the reliability or security of Internet (even if improvements are currently sought), but to the need for sharing information. Frequent problems that arise are due to non common language or format, multiple point-to-point interactions, static inflexible organisation of information, multiple, disparate catalogues, and case-to-case translators, etc. The problem is particularly acute when a large number of trading partners attempt to agree and define the standards for interoperation.

The ontology approach focuses on solving these problems by creating consistent behaviours among participants and allowing complex interactions and a flexible environment, which do not require all participants to adopt common infrastructure. But for enabling commerce agents to interact without misunderstanding, the commerce agents must adopt common ontologies. These ontologies must be developed and shared among agents.

In that project, the role of ontologies will be to facilitate man-machine dialogues between user and software agents. Ontologies form a knowledge base for agents to infer information that is relevant to the user's request and they will be used for information filtering, classifying and indexing.

We plan to exploit existing ontology models as a consensual representation of a field allowing the exchanges independently of the language of the end-user, the service or the content providers (CP). Various ontologies are required and play a central role in *i)* mapping the human language query onto an ontological representation; *ii)* producing composite services from CP catalogues by the broker; *iii)* mapping ontological queries onto the ontological representations of the catalogues, and *iv)* describing the domain knowledge. Building a domain ontology requires a deep understanding of the domain, thus we will reuse a maximum of existing ontologies available on Ontology Servers: (Ontolingua Server[16], Ontosaurus [17], WebOnto [18], etc.). This implies to find a technique for evaluating ontologies and also for inter-ontology management. This will be our first challenge.

Moreover, ontologies will be used for classifying, filtering and referencing the external information to be processed. This will be taken care by several ontology servers: *i)* Domain ontology servers, *ii)* E-commerce ontology server, *iii)* Third party negotiation ontology server, *iv)* Ontological CP catalogue. Due to market dynamics, and the evolution of the catalogue contents and the CPs, the ontological representations will be made adaptive and evolving. Our main focus is on the domain ontology for the CPs, for whom two basic services are provided: multilingual access and automatic cataloguing and multilingual / cultural customised presentation.

Ontologies will be also used for improving the performance of traditional information retrieval system [1] [2], [3], [4], [6], [7]. Most of the current browsing and navigation tools use keyword-based search, and when the requested keyword does not match the indexed one of the relevant document, the system does not find it. The new approach is to use ontologies [10] for cross-lingual information retrieval systems. The multilingual approach, which is the second main challenge of that project, will be handled more thoroughly in our later papers (e.g. [9] summarises available solutions). Other approaches like SHOE [19] and Ontobroker [15] perform ontology-based information retrieval, but they are not multilingual. EuroWordNet [12] is a step in that direction, but is not yet applied to information retrieval. In our project, the ontological representation of the CP catalogues and the ontological queries will be made as language-independent as possible. From the information retrieval aspect, ontologies will improve the accuracy in fuzzy information search. Viewed from natural language processing, ontologies will facilitate mono- and multi-lingual human-computer dialogues by paraphrasing the query of the user through context identification and disambiguation.

Finally, the use of ontology aims also to solve the cross-language problem as there is not always 1-to-1 correspondence of terms between different cultures. The system should be able to find the same product even if the query is made in different languages to that of the product name / attributes. This possibility should be as natural as possible to the user.

2.2 Other components

The data communication layer and electronic commerce architecture will be implemented by making use as much as possible of available components (typically from 4th framework ACTS projects [20] where MKBEEM participants have relevant and exploitable results) and by using existing platforms exploited by consortium partners (e.g. E-Commerce platforms).

The existing application layer, on the contrary, will be enriched with *innovative service components providing multilingual interface for both the end-user and the Content/Service Provider, judicial and electronic commerce ontologies and other related Domains Knowledge, supply-demand matching facilities, and support for intelligent contractual transaction management*.

In order to demonstrate our results, a full end-user service prototype utilising the mentioned component services will be implemented on a pan European scale. This will be achieved in two iterative steps: one in Month 13 with basic end-to-end functionalities demonstrated on a limited geographic scale, and one in Month 26 with the full specification on three European sites (Finland, UK and France).

3 A MULTILINGUAL MARKETPLACE

Two Prototypes of the Mkbeem services are planed. The first one is due March 2001, and the final one April 2002.
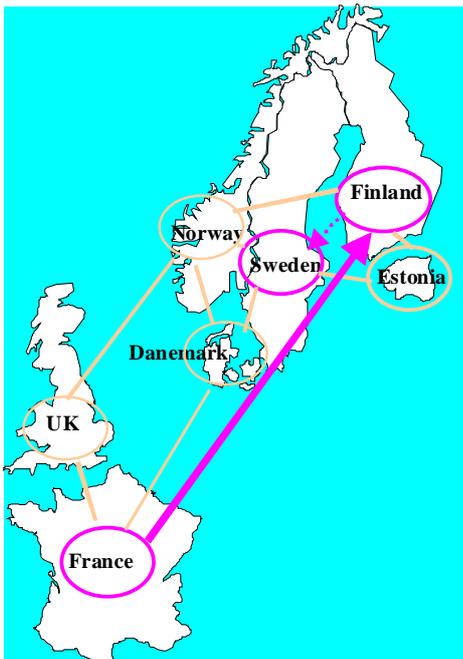
Knowledge and Technology components for industry to build on for exploitation in an unforeseen yet and wider scope.

The project envisions the following important benefits:

Drastic extension of the current electronic commerce platforms to a broader spectrum of products and services on a pan-European basis;

- Natural language interface for both the end-user and the Content Provider / Service Provider ;

- Automatic multilingual cataloguing for the Products and Service Providers ;

- On-line e-commerce contractual negotiation mechanisms in one's language that guarantee safety and freedom ;

- One stop shopping paradigm "European Market Square" ;

- Transparency of linguistic and cultural differences ;

- Human language expression for the development of individual and collective ability of using human languages and so to protect cultural diversities.

The main expected results can be summarised in the following tables:

Technology Results
• NLP in three core languages: English, Finnish, French,
• Plus further demonstrations for: Swedish, Spanish
• Ontology for 4 domains
• Agent based architecture

Trials results
• Multilingual services in three complementary sectors
• Validation of generic architecture
• Validity and Exploitability of the architecture: NLP and Ontologies
• Multilingual content classification & Management
• Juridical reconciliation by Ontology

In the standardisation field, as active members of the FIPA consortium (Foundation for Intelligent Physical Agents), and as the FIPA'97 and FIPA'98 specifications are available [14], as declared recently to FIPA, we will try our best to use - whenever it is justified - the FIPA specifications. In particular, the Speech Acts (FIPA97 - part 2) and the Ontology services (FIPA98 - part 12) published in fall 1998.

In the Knowledge Engineering area, active contributions to the scientific community and pre-standardisation are already foreseen.

4 PERSPECTIVES

The MKBEEM consortium being composed of Basic Technology providers, Systems integrators and Users, we will endeavour to produce useful services for all the citizens and re-usable pieces of

5 REFERENCES

[1] Alon Y. Levy et al., Query answering algorithms for information agents, *AAAI'96 Proceedings,* 1996.

[2] Sudarshan Chawathe et al., The TSIMMIS project, *IPSJ Proceedings*, Tokyo, Oct. 1994.

[3] Yigal Arens et al., Query reformulation for dynamic information integration. *JIIS*, N°6, 1996.

[4] Michael Genesereth et al., Infomaster, *ACM SIGMOD-97*, 1997.

[5] Marie-Christine Rousset et al., The PICSEL project, *ECAI'98 Proceedings*, 1998.

[6] LRI-FT R&D, The PICSEL project web site : http://www.lri.fr/~picsel/

[7] Alon Levy, Marie-Christine Rousset, CARIN, *AI journal*, 101, 1998.

[8] L. Redmond. The virtues of controlled language. *Le Journal: http://www.linglink.lu/LeJournal/*

[9] Aarno Lehtola, Multilingual Information Services as a Goal. *HCI International `99*, 1999, pp. 767-771.

[10] Kuldar Taveter et.al., Ontology-based Query Translation for Legislative Information Retrieval, *UI4ALL 99, GMD Report 74*, 1999. pp. 47 - 58.

[11] G. Erbach et al., MULINEX. *Proc. AAAI Spring Symposium*, Standford Univ., 24-26 March 1997.

[12] EuroWordNet *http://www.let.uva.nl/~ewn/docs.htm*

[13] The role of formal ontology in the information technology, Int. J. Human-Computer Studies, Vol.43 (1995).

[14] FIPA - "FIPA'97 Specification Foundation for Intelligent Physical Agents" – drogo.cselt.stet.it.fipa – 0ctober 1997

[15] Fensel, D, Decker, S. Erdman M. Studer, R. Ontobroker: The Very High Idea. In Proceedings of the 11th International Flairs Conference (FLAIRS-98), Sanibal Island, Florida, May 1998.

[16] Farquhar, A., Fikes, R., Rice, J. The Ontolingua Server: A Tool for Collaborative Ontology Construction. Proceedings of KAW96. Banff, Canada,1996.

[17] B. Swartout, R. Patil, K. Knight and T. Russ. Towards Distributed Use of Large-Scale Ontologies. Spring Symposium Series on Ontological Engineering. Stanford University, CA. 1997, Pages: 138—148.

[18] Domingue, J. Tadzebao and Webonto: Discussing, Browsing and editing ontologies on the web. In Proceedings of the Eleventh Knowledge Acquisition Workshop, KAW98, Banff, 1998.

[19] Luke S., Heflin J. SHOE 1.01. Proposed Specification. SHOE Project. February,2000.
http://www.cs.umd.edu/projects/plus/SHOE/spec1.01.htm

[20] ACTS-ABROSE, ACTS-ABS project
http://b5www.berkom.de/ABROSE, http://b5www.berkom.de/ABS/