LET'S SEMANTICISE THE WORLD!! … OR NOT??

Michele Missikoff
LEKS - Lab for Enterprise Knowledge and Systems, IASI-CNR, Italy

Keywords: Knowledge Economy, Ontology, Semantic Annotation, Semantic Business Process, Semantic Business Services, Mapping Discovery, Enterprise Model, Model Driven Architecture.

Abstract: Many say that Knowledge is the oil of the Third Millennium. The European Council, in the Year 2000, launched the Lisbon Strategy[1], aiming at building in Europe the major Knowledge Economy in the world. Then we had the 9/11 and the "subprimes" (the future is rarely as we expect) and an epochal economic slowdown, without an equivalent since 80 years. Nevertheless, the knowledge economy, pushed by the consistent evolution of the information technologies, is steadily progressing and further expansion is foreseeable in the near future. We believe that for a renewed sustainable growth, semantic technologies will play an important role. In this paper we briefly draw the main lines of a possible future evolution of the application of semantic technologies to the business world.

1 INTRODUCTION

The knowledge moves the World. The human beings are not among the fastest or the strongest animals on the Planet, nor they are protected by a special fur or by a hard carapace. However, they acquired the leadership on the Planet thanks to the knowledge, and the capacity to use it for practical purposes.

With the advent of ICT there has been a tremendous impulse on how humankind deals with knowledge. Since the first wave of computers, in the social and industrial realities of the 60s, the knowledge managed by a computer can be seen divided in two different sorts: knowledge for computers, e.g., software programs, and knowledge for humans, e.g., digital representation of paper documents. In the former case, the computer is able to execute the software, without "understanding" (this term will be better clarified later) it, while in the second case the computer is just a container (i.e., incapable of execution nor understanding). In parallel, Artificial Intelligence, and in particular Knowledge Representation methods and systems, started to develop solutions, including languages and reasoners, to provide the computer with some forms of "understanding" and execution of knowledge. This represents a third sort of knowledge, somehow positioned between the two above.

To better clarify the different perspectives, human- and computer-oriented, with respect to knowledge understanding and processing, at an intuitive level we may draw a diagram where on the x-axis we place the three knowledge environments corresponding to the three different sorts of knowledge: (i) document and content management systems (CMS), (ii) semantics representation and processing systems (SRS), (iii) programming and execution systems (PES). Then we draw three lines. One concerning the ease of access for humans to the knowledge represented in the three environments, the second represents how easy is for the computer to execute (or interpret, if you prefer) the knowledge, and the third how much the computer "understands" the knowledge[2] therein.

Time passes, and the basic issues of programming and content management did not change much. Conversely, knowledge and semantics representation methods and tools have evolved, thanks to the energies spent in research and development, with the objective, among others, to achieve a unified knowledge space: same knowledge (possibly, in different digital formats) for humans and computers, and similar "understanding" and pro-

[1] http://en.wikipedia.org/wiki/Lisbon_Strategy, visited on April 2009

[2] For "understanding" we intend here the capacity of an active entity to acquire knowledge and be able to modify its behavior on the bases of to the acquired knowledge (we disregard here the pure speculative knowledge).

Matsumoto M.
SERVICE COMPUTING EIS, WORLD PANIC AND OUR ROLE CHANGE.
DOI: 10.5220/0006811800010001
In *Proceedings of the 11th International Conference on Enterprise Information Systems (ICEIS 2009)*, pages 5-15
ISBN: 978-989-8111-84-5

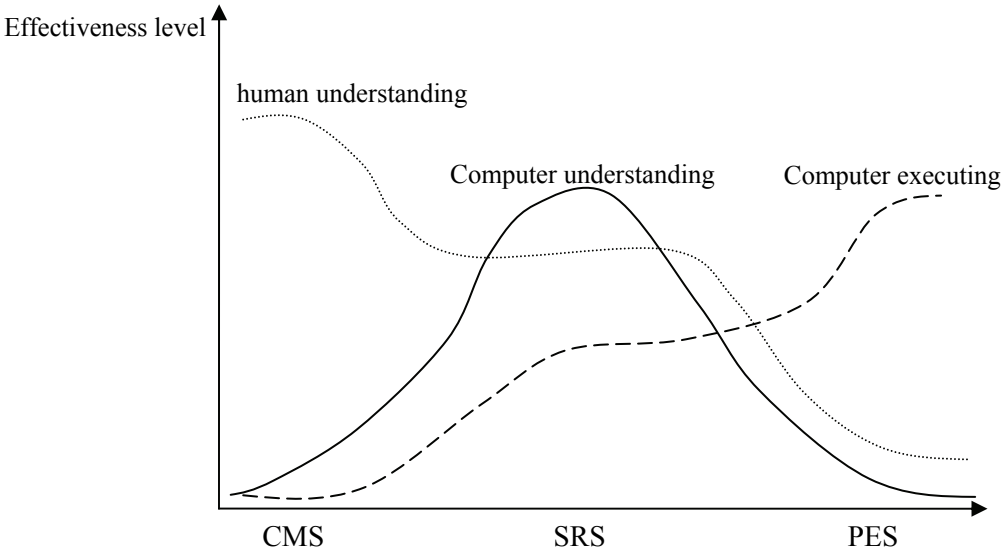


Figure 1: The KM effectiveness.

cessing capability. Today, the "Mission: Impossible" is the progressive convergence towards such a unified knowledge space.

May be, the Impossible Mission becomes more possible if we restrict our world to enterprises. Today, an enterprise produces every day an incredible amount of documents, all of them in digital form (from technical reports to meeting minutes, from emails to field studies, to market analysis). Therefore, we have a sort of Enterprise Image in Digital Form (EIDiF) that reflects all the possible knowledge, produced and/or acquired, that traverses across the enterprise organization, the production, the marketing, and all the other departments of the company. The EIDiF primarily consists in the whole set of human-oriented knowledge (i.e., digital documents, but comprises also the vast amount of information typically maintained in the enterprise databases). The content of the documents, as they are produced and exchanged by humans, cannot be directly understood by computers. To allow computers to access such a content, there is the need of a significant pre-processing: the semantic content of the documents need to be extracted and represented in a formal way. This is one of the primary goals of semantic technologies.

In recent years, semantic technologies are proving to play a central role. New (and renewed) solutions are spreading, by using formal knowledge coding, ontologies, reasoners, and semantic annotations, addressing both the static (e.g., business objects) and the dynamic part (e.g., business processes) of the enterprise. We are witnessing the progressive evolution of such technologies towards the achievement of the aforementioned Unified Semantic Space. Indeed, we see that computers and humans are increasingly cooperating, helping each other in more sophisticated, knowledge-intensive activities (naturally, having humans superordinated, to avoid, e.g., an Asimov scenario).

As anticipated, a key issue in this scenario is the possibility of (semi)automatically extracting the semantic content of enterprise documents, encoding it in a machine "understandable" form. In essence, it means to develop a formal semantic theory of the addressed matter, of the enterprise and the context it operates in. This is today a clear trend, but how much will it be possible to capture and model in formal terms? There are various signs showing that we are indeed proceeding along this line. From enterprise ontologies, to formal business process modelling, from automatic knowledge extraction to business rules, the research is eagerly progressing in the direction of developing new formal methods to model and manage the business reality. Furthermore, there are important theories that can be applied in specific business sectors, such as chaos theory and system dynamics (not to mention applied mathematical theories that go from finances to logistics, from market analysis to accounting).

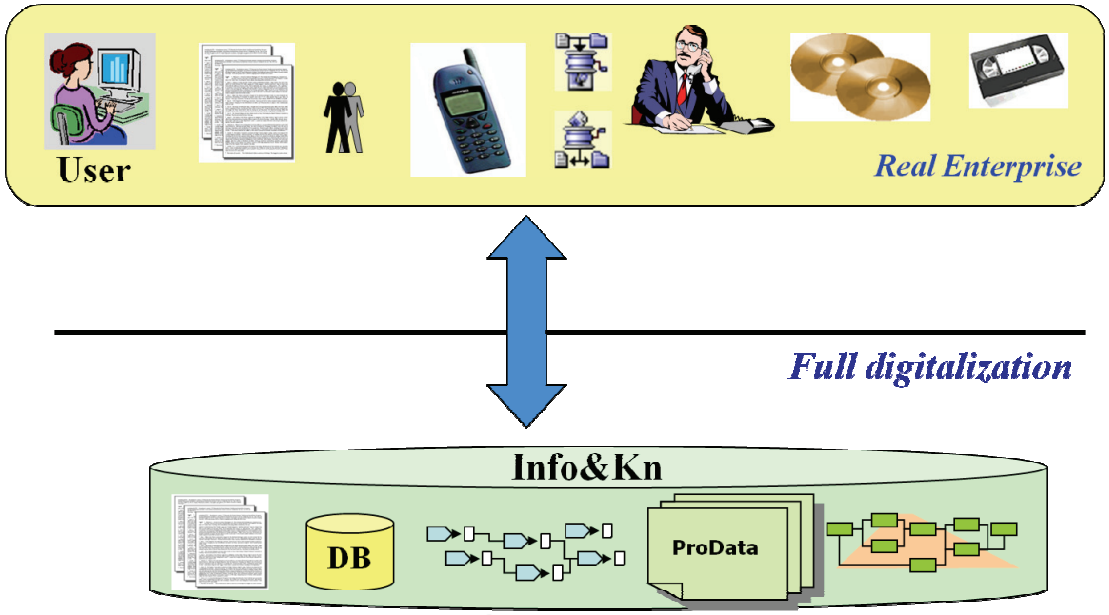


Figure 2: The Enterprise Digital Image.

The target is a business world where a rich collection of formal theories and precise models allow the computer to meaningfully manage an increasingly large fraction of the business knowledge, covering the large majority of activities, behaviours, objects, and actors that are involved in business transactions.

In the current scenario the research and innovation is proceeding in a very fragmented way. New theories produced by different disciplines, may be effective on specific problems, but are hard to relate each other to achieve a coherent, comprehensive solution. We still have a huge level of fragmentation that, optimistically, yields to a sort of federation of different theories, produced by different communities, with very limited capacity of integration.

But reality is one, so the goal is the construction of a holistic formal enterprise theory. There have been scientists who anticipated a vision of this sort. For instance, the work of Max Tegmark (Tegmark, 2008) whose visionary approach can be condensed in the idea of a mathematical TOE (Theory Of Everything) based on the notion of a Mathematical Universe (MUH: Mathematical Universe Hypothesis). While Tegmark is trying to build such a vision in a top-down approach, there are several other promising research lines that proceed bottom up. It is worth citing the emerging field of Econophysics (Rosser, 2006). Here the researchers acquire all possible quantitative data representing any possible business phenomena. Then, advanced rigorous methods, mainly drawn from theoretical physics, are applied to discover regularities. Very promising results are obtained, for instance, by applying statistical quantum mechanics.

Any mathematical theory, to be fully effective and in order to relate a mathematical theory to the real world (the enterprise world, in our case), needs to be associated with a clear semantics. The latter has the essential goal of creating a bridge between symbols and expressions in a mathematical theory and objects and events in the real world. In essence, a mathematical theory always needs a semantic apparatus connected to it, no matter how advanced, sophisticated, accurate it is, to explains and predict relevant phenomena in the modelled (business) domain. Such a semantic apparatus can assume different forms. It can be mainly intuitive (for instance, simply using natural language, with words connected to mathematical expressions, variable symbols, quantities, and other elements of the theory) or it can be a formal theory on its own part.

The idea is that a comprehensive, encompassing, accurate, effective formal semantic infrastructure, conceived for a given business sector, will be capable of correctly positioning and cross-relating all the specific theories of different sorts that are proposed for modelling and explaining different sectors and viewpoints in an enterprise. For this reason, we will proceed in analysing the possibility of building such a semantic infrastructure of an enterprise based on a formal setting, achieving what we refer to as the Semantic Enterprise.

2 TOWARDS THE SEMANTIC ENTERPRISE

In proceeding in the direction of the semantic enterprise, it is useful to first sketch a layered architecture where the various semantic services can be placed in a rational way. We start to illustrate the sketchy architecture reported in Figure 3 starting from the bottom layer.

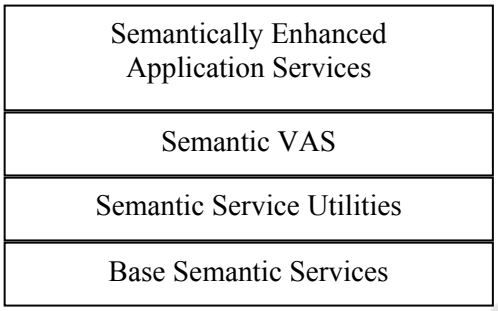


Figure 3: A semantics driven architecture.

2.1 Basic Semantic Services (BSS)

We start bottom up, with the aim of building a supporting semantic infrastructure. Such an infrastructure comprises three essential elements, and the corresponding platforms and services.

Ontologies. When we talk about semantics, we primarily intend, in accordance with (Ushold 1998), a set of concepts with their relationships, internal structure (e.g., represented by the associated properties), and constraints (i.e., axioms to be enforced). An ontology (a taxonomy of structured concepts) can be formally represented by using some sort of logic.
Given a complex reality, such as an enterprise or a business domain, generally it is practical to develop more than one ontology, each of which having a clear focus. E.g., a marketing ontology, production ontology, HR ontology. But the reality is one, and humans tend to segment it to be able to cope with the complexity, therefore the segmented ontologies need to be related each other. This can be achieved with an inter-ontology mapping infrastructure.

Semantic Annotation. The rich and articulated collection of documents and (factual) data sets that we can find in an enterprise always refer to real business entities, representing therefore the EIDiF (i.e., the digital image of the enterprise): the products, the people, the markets, the competitors, etc., where each element class has its own form of digital representation, typically, to be presented to and managed by humans.
A key step in the "semanticisation" of the reality is the systematic creation of mappings between entities and phenomena of the business reality and concepts of one (or more) ontology. We refer to such mappings as *semantic annotations*. In its simplest form, a semantic annotation is a direct link between a concept in an ontology and an element of the enterprise. E.g.,: (Mario ←→ Employee). In general, to provide an articulated account of a fragment of the business reality, it is necessary to create an ontology expression.

Reasoner & Truth Maintenance. The two above structures, ontologies and semantic annotations, are essentially repositories of knowledge. Then we need to complete this layer with an engine capable of processing the above structures to perform two basic computations: (i) derivation of new knowledge and (ii) verification that the existing structures do not contain contradictions.
These functionalities are necessary for the very same management and maintenance of the semantic repositories. In fact, the reality continuously evolves and the enterprise documents will reflect every significant evolution, therefore the semantic repositories, that reflect the state of the affaires, must evolve accordingly. If we add a new piece of semantic knowledge in an ontology that we know is free of contradiction, a TMS (truth maintenance system) is responsible to guarantee that also the new version of the ontology will be released in a consistent state, i.e., we did not introduce any elements that invalidates what we already know[3].

2.2 Semantic Service Utilities (SSU)

On top of the base semantic services we can build a number of semantic service utilities. A service utility is a software facility available to everybody, accessible with a predefined universal protocol at predefined published conditions (cost, SLA, etc.). The proposed SSUs provide the key functions on top of which any other possible semantically-enhanced application can be built. In turn, to exist the SSUs need the BSSs seen in the previous paragraph. The key SSUs are listed below.

[3] Here we touch a first limitation of the current semantic technology: the limited capacity of representing contradictions. While we know that the reality is full of contradictions

Semantic Knowledge Mining. This service utility is capable of accepting a human-oriented document (from a plain text document to an XML file, from a business process diagram to an email) and an ontology, returning a set of semantic annotations. In essence, this service utility is capable of extracting the semantic content of a document, building an explicit, formal ontology-based structure (therefore, by using different ontologies it is possible to have semantic annotations emerging from different disciplines, different perspectives).

Semantic Matchmaking (SMM). This service utility is based on some *mapping discovery* methods. Mapping discovery is a very vast research area aimed at solutions that, given two structures, are able to derive a formal relationship that characterises the correlation between the two structures as a whole and among the individual elements of the two structures (in case, considering also relevant sub-structures). There is a wealth of available methods in the literature, with a great variety of structures considered (from graph stractures to logical theories, from geometrical figures to natural language sentences and paragraphs). Here we restrict our focus to ontologies (or fragment of them). The output of a SMM service utility can be of various sorts from an algebraic condition, such as equivalence, containment, of disjunction, to a quantitative measure (e.g., a value between 0, in case of disjunction, and 1, in case of identity).

Semantic Interoperability Service Utility (SISU). This service utility is necessary when we have two structures and, once we identified with the previous SMM that there are divergencies, we intend to reconcile such differences. This service is similar to the previous SMM, but while in the previous case we intend to identify the similarity degree, as a static declarative parameter, here we intend to identify an active mapping, that is the operations necessary to transform one structure into the other. The SISU is capable of transforming the messages exchanged between different software applications that interoperate by exchanging information, therefore understanding each other messages despite the difference in their respective data organizations.

2.3 Semantic Value Added Services (SVAS)

This third layer of semantic services is still of a general nature, but the offered facilities are able to concretely contribute in producing value for the enterprises. We briefly list some of them.

- *Semantic Search and Retrieval*. This will be the new frontier of search engines. The injection of semantics in search engines will significantly improve the performances, effectiveness, and the user satisfaction. Especially when developed jointly with user profiles (see the next point).
- *Semantic User Profiling and User-centred HCI (Human-Computer Interaction)*. User profiling is a very promising area. But the developed solutions will be even more effective, if the user profile will be enriched with semantic annotations. In particular, when a semantic user profile will be used for search purposes, the semantic search engine will be able to consider it, jointly with, for instance, contextual information, to rewrite the user query and to optimize its execution.
- *Staffing and Experts Team Building*. The semantic profiles (including competences and skills) can be used for the optimal composition of working groups, where the gathered competencies and skills are suitably blended with respect to the activities to be performed.
- *Enterprise Consortium Building*. Here we change scale, moving up to the level of a consortium of enterprises. When a consortium is built to respond to a business opportunity (e.g., a public call for tender) it is necessary that the gathered enterprises show a good coverage of the capabilities required in the call. A joint semantic analysis of the Call and the enterprise profiles will provide important elements to proceed in the formation of the consortium.

2.4 Semantically Enhanced Applications Services (SEAS)

Here we address specific (vertical) enterprise applications, such as Accounting, HR Management, Production Planning, Sales and Distribution. Semantic technologies can have a wide potential impact, empowering all possible enterprise applications. Some enterprise applications will deeply change with the injection of semantic capabilities, but some other will simply disappear to be substituted by new integrated Business-IT solutions, not conceivable without the use of the semantics. The innovative solutions will emerge from the joint use of semantics and Web 2.0, user-

centred social software, significantly impacting different industrial sectors. Among the most innovative vertical applications, we may cite:

- Intelligent autonomic logistics systems
- Disaster and emergency prevention and management
- Advanced cross sectorial health care
- Cognitive economics

3 THE KNOWLEDGE ENTERPRISE NETWORK

The full achievement of the knowledge enterprise, based on the semantic technologies, cannot be reached with the enterprises we know today.

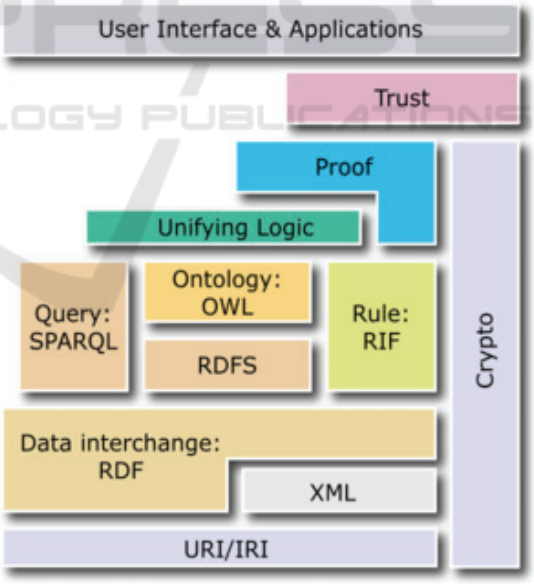
3.1 In Search for New Enterprise Models

In parallel to the technological innovation, it is necessary that the organization and operational models of the enterprise undergo a deep change as a precondition for the full deployment of the potential of the knowledge infrastructures. In this perspective, we may recall the vision of Stafford Beer, rephrasing it in the perspective of the semantic enterprise: asking how to use semantic technologies in the enterprise is a wrong question, the right question is what will be the transformations that semantic technologies will induce on the enterprises. But, even more correct is to ask what will be the enterprise of the future once that semantic technologies will be fully deployed. Then, let's try to depict the main lines of such a possible future.

In the previous section we presented a framework for the development of enterprise wide software applications based on the extensive use of semantic services. In fact, Figure 2 represents the layers of a Semantic SOA that will be realised in the next decade or so. In this perspective, the semantics is mainly involved in the achievement of the advanced software services, but not in the production of the available services: there is still a significant amount of software to be developed, debugged, tested, and maintained.

3.2 From Software Programming to Knowledge Representation

Pushing further our vision, we can envisage the future knowledge enterprise network, where new paradigms for service development will be totally based on the enterprise semantics. The resulting services will not be coded with the traditional software techniques. Enterprise IT applications will be characterised by a separation between business logic and business operations, pushing the MDA (Model Driven Architecture) to its full accomplishment. Essentially, the business logic, including strategies, rules, and high level best practices, will be represented by Semantic Business Processes (SemBP), while activities and operations will be represented by Semantic Business Services (SemBS). The latter, if necessary, will be recursively expanded, showing the internal structure in the form of more detailed SemBPs, while the components will be lower level SemBSs, until atomic SemBSs will be reached. The latter are SemBSs that can be fully (and operationally) specified in one unambiguous step (human or automatic), to be executed either directly in the enterprise or by an external entity.
SemBP and SemBS will be (semi)automatically derived from the extensive repositories of enterprise documents. They will be represented in a declarative form, e.g., rule-based, with the support of reference ontologies.



The supporting networked infrastructure will be the evolution towards a semantic version of the FInES (Future Internet Enterprise Systems) that we know today[4]. A possible global architecture has been suggested by W3C, with the Semantic Web "cake".

[4] http://cordis.europa.eu/fp7/ict/enet/ei_en.html

3.3 The Unified Enterprise Knowledge Space

The key idea of the scenario reported above is a tight integration of human and computer knowledge, with synergic capabilities of understanding and proactively using the available shared knowledge: a symbiosis between natural and artificial intelligence. We assume that the principles, rules, and operations, according to which an enterprise functions and produces value, are all represented in a form or another in human-oriented documents. From business processes to roles and positions of the personnel, from marketing strategies to assets management, more and more there is a consistent (or supposed so) production of strategic, tactical, or operational documents. We can imagine that the semantic technologies will be able to extract the knowledge therein reported and codify it in the forms that can be interpreted by a computer system. In essence, when the management of an enterprise decides to introduce a change in the operations of the organization, this intention is reported in one or more documents. These documents will be transmitted to the interested sectors and people therein, who will modify organizations, operations, and employees behaviour accordingly. But in parallel, there is a need of changing the enterprise application systems for the parts affected by the above mentioned decisions. In the knowledge enterprise, the management documents will be semantically analysed, the knowledge mining services will extract the new instructions to be propagated to the operational knowledge repositories (i.e., those containing the SemBP and the SemBS). Semantic Matchmaking services will discover where (i.e., on which processes and services) the new directions will impact. Then, the updated operational knowledge will be activated, to guarantee that the IT enterprise applications will behave according to the new directions (in this way, the well known Business/IT alignment will be largely solved).

4 THE OBSTACLES TO THE ADVENT OF *KEN*

The progress towards the new Knowledge-centric Enterprise Network (KEN) models, made possible by the extensive use of knowledge and semantic technologies, has been sought since long time. More than a decade ago there have been the first important results in this direction (see for instance (Fox et al., 1998) and (Uschold et al., 1998). Similarly, the Semantic Web has been proposed at the beginning of this decade. The decade of the 90s has been characterised by the explosion of the Web, the current decade is characterised by the Social Web. Many say that the upcoming decade will be finally that of the Semantic Web and, as its natural consequence on the business world, it will see the advent of the Knowledge Enterprise Network. We know that the technological innovation is unstoppable, however there are several factors that may delay the joint evolution of the technologies and the enterprises along the lines described above. Here we can briefly summarise a few of the hindering factors. Some of them are real, some other depend on the wariness of the majority of the businesses.

Knowledge Culture. We need that the enterprises develop a diffused culture of semantics, a better awareness of the advantages that a diffusion of such technologies will bring to enterprises, and to the society at large.

Semantic technologies are still considered as research artefacts, far from the practical and convenient level of maturity, in order to be actually adopted by enterprises.

Semantic technologies are very demanding in terms of processing power and storage systems. We still need a great deal of research before their extensive application in real industrial settings. To adopt semantic solutions, and enterprise needs a group of highly specialised experts, such as ontology engineers or logicians.

Last but not least, costs and risks still appear high with respect to the expected (but not concretely proved) benefits.

5 CONCLUSIONS

In this paper we sketchily presented the main lines of a possible future, where new technological solutions will concur in the realization of a unified semantic space. In this space humans and computers will have access to the same knowledge, being therefore able to tightly cooperate in different high level activities. In such a reality, there will be a parallel evolution of the business dimension, new enterprise models will emerge and the notion of ***value*** will be different from our understanding of today.

We are at the verge of a new decade witnessing an acceleration of the progress towards the Knowledge

Economy. But we know that the progress hardly follows a linear trajectory, and there is no positive determinism in the technological evolution. Therefore, as usual, we are the creators of our own fortune.

ACKNOWLEDGEMENTS

I wish to thank the colleagues Antonio De Nicola e Francesco Taglino who helped me to better focus the vision reported in this paper. Furthermore, I wish to thank Sergio Gusmeroli for the long and open discussions we had in digging the frame and mission of the COIN European Project, that influenced the layered service-oriented vision reported in Section 2.

REFERENCES

D. Bianchini, S. Castano, F. D'Antonio, V. De Antonellis, M. Harzallah, M. Missikoff, S. Montanelli; Digital resource discovery: semantic annotation and matchmaking techniques, In proceedings of Interop-ESA06 Conference, Bordeaux, March 2006.

M. Fox, M. Gruninger; Enterprise Modeling. AI Magazine, V. 19 n. 3, AAAI Press, 1998.

Fabriani, M. Missikoff, P. Velardi: "Text Mining Techniques to Automatically enrich a Domain Ontology", Applied Artificial Intelligence Journal, Vol. 18, No. 3, May-June 2003.

T. Berners-Lee, J, Hendler and O. Lassila; The Semantic Web. Scientific American, May 2001.

J. Barkley Rosser; Econophysics. In The New Palgrave Dictionary of Economics, 2nd edition, January 2006.

Nigel Shadbolt, Wendy Hall, Tim Berners-Lee; The Semantic Web Revisited. IEEE Intelligent Systems, May-June 2006.

M. Tegmark; The Mathematical Universe. Foundations of Physics, V.38 n.2, February 2008.

M. Uschold, M. King, S. Moralee, Y. Zorgios; The Enterprise Ontology. The knowledge engineering review, Cambridge Univ Press, 1998

E. Von Schweber; Roadmap for Semantics in Netcentric Enterprise Architecture. US General Services Administration Report, February 2006.

BRIEF BIOGRAPHY

Michele Missikoff, Director of Research at IASI (Istituto di Analisi dei Sistemi ed Informatica) of National Research Council, and Coordinator of LEKS: Lab for Enterprise Knowledge and Systems since 1999. Scientific Director of the European laboratory on Enterprise Interoperability, Interop-VLab, has a long experience in European projects, both in eBusiness and eGovernment. Elected member of the National Steering Committee for Information Science and Technology, from 1992 to 1997. He has been member of the editorial Board of VLDB Journal and co-founder of EDBT Conference series. Past-president of EDBT Endowment, he has organised numerous conferences and workshops. He has pubblished more than 150 technical and scientific papers, the majority of which on international journals and conference proceedings.