

Competence Ontology for Network Building

Kafil Hajlaoui , Xavier Boucher, Michel Beigbeder, Jean Jacques Girardot
Ecole Nationale Supérieure des Mines de Saint Etienne G2I,
42023 Saint Etienne, France
{hajlaoui, boucher, mbeig, girardot}@emse.fr

Abstract. The overall objective of this research is to engineer a decision aid approach to support the identification of collaborative networks which could then constitute potential Virtual Breeding Environments. An interesting aspect of this approach is working with internet and thus considered an "open universe" of potential partners. In this perspective, the paper puts the focus on extracting essential facets of firm competences using an ontology approach. The method followed to construct the ontology is presented as well as a brief introduction on its use. This work is part of an on going project to produce cognitive aids to support decisions when seeking to form partnerships and establish VBEs.

Keywords: Information extraction, Ontology, Competence model, Enterprise networks.

1 Introduction

A lot of research work deals with the formalization of characteristic data concerning potential partners for networked organizations [3][13][4]. Most of these approaches are adapted to a semi-closed environment defined by a Virtual Breeding Environment (VBE). VBEs provide a collaborative environment, facilitating trustable exchanges of information to help the selection of partners when creating virtual organizations.

The research we refer to focuses on a complementary step: furnishing a decision aid support for identifying potential VBE. This is based on the hypothesis of an open environment of potential partners to build VBEs where any company can participate. Typically, this issue often appears when you have to analyse a regional business area to identify potential collaborative networks among firms. As a consequence, the identification of potential collaborative partners will be based on the use of public information, available through the public web sites of the companies. This assumption induces the use of specific information extraction mechanisms.

As in the coordination approach the information extraction procedure focuses on two key factors: the activity fields of the companies and their internal competences (see justification in [9]). This information on company activities and competences will be further used, at a second stage, to generate new knowledge on the potential structure of VBEs. The overall approach of such a decision aid has been already described in [10].

This communication concentrates on a competence ontology required to extract synthesized information on company proficiencies. Due to the aforementioned complexity of this competence concept, a semantic oriented approach is required for the extraction. Ontology is used as a semantic resource to guide the extraction process applied to competences. One of the objectives of this paper is to explain the method used to engineer the ontology. Furthermore, its use, with syntactic and semantic patterns which support the information extraction procedure, is underlined. These patterns can treat semantically and pragmatically the data available on company web sites. To reduce the ontological complexity, our research is limited to a sole enterprise activity domain: the mechanical industry.

In section 2, some essential references on competence modelling are provided as well as the basic competence model used as the starting point of the ontology. In section 3 the methodology utilized is described along with some ontological components. In section 4, we briefly introduce the use of this ontology for information extraction. Finally, our conclusions are presented.

2 Competence Modeling

The competence management is an important lever for business competitiveness and for cooperation between firms. With an increasing production after the 90's, scientific literature published quite a large amount of research examining how to better characterise the notion of competence. Several competence modelling approaches have been suggested, for the implementation of methods and tools for competence management. Table 1 provides some examples of contributions from three different scientific communities, which intend to manipulate the notion of competence using more or less formalized models. This table is not exhaustive, but already underlines the diversity of publications on this modeling issue.

Referring to a larger state of the art developed in our research, we present in figure 1 a basic model of competence used as a root to develop a more detailed competence ontology (see section 3). This model refers to the key notions of the Resources Based View for firms: the competences of a company emerge as a combination of internal capabilities. These capabilities are the result of the activation of different resources acquired by the company. To define further these key notions needed, we consider four basic types of resources: human, technological, informational, organisational resources. Furthermore, two basic types of firm capabilities are distinguished: methodological linked to the added value provided by the working methods used by a company to deliver its products or services; and technological referring to the creation of added value based on the use of technical resources and processes. This structure supplies a generic and abstract model used as a starting point to create an ontology.

Table 1. Examples of references on competence models, from diverse sectors.

Ref	field	Key points
[1]	Industrial management	CRAI model (Competence Resource Aspect Individual) for enterprise modelling. This semantic model represents the links between competence and other enterprise modelling constructs: the context, resources, individual and activity missions.
[12]	Industrial management	A formal model of competence is proposed to be integrated within an organization model for the management and development of competences. This supports a quantitative approach of competence assignment, applicable both to individual and collective competence
[4]	Computer Science	In the context of VO creation the authors present a generic model of competence based on the concepts of Capability, Capacity, and Conspicuity. This was created for the management of firm competence profiles, so as to make possible evaluation of competence profiles.
[11]	Computer Science	Firm competences are described by a set of qualifications, technologies and knowledge, but without referring to a more generic model of competence. Based on the use of a Multi-agent System, some user scenario are proposed to create business collaboration among companies, with an explicit view on their internal competences.
[14]	Management Science	An ontology formalises the concepts necessary to represent and manage the corporate memory of a company. This ontology is suggested as a support for competence management.
[15]	Management Science	The Resource Based View of the firm initially proposed by Wernerfelt has generated a large amount of scientific work. His theory provides notably a stable definition and articulation of the crucial concepts necessary to describe firm competences.

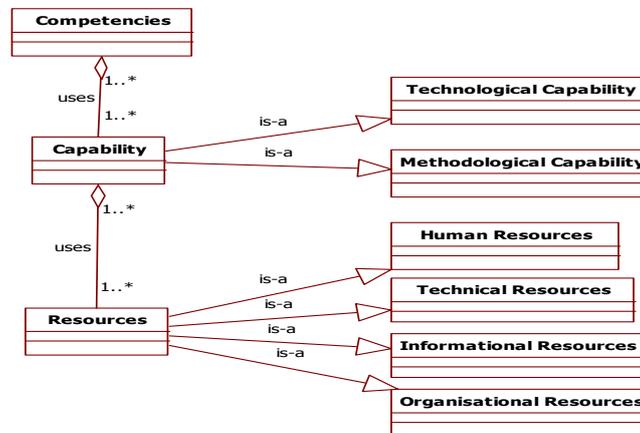


Fig. 1. Basic competence model

3 Ontology Engineering

As mentioned in section one, the final information sought is a similarity indicia among the overall competences of individual companies. In that objective, our

approach consists in extracting a set of “traces” concerning company competences. Consequently, the ontology required here is that of “competence traces”, dealing with only sector of activity of the mechanical industry, as mentioned before. Of course, on a company web site there are a lot of separate phrases, referring to a large variety of concepts (human or technological resources, activities, processes, working methods...) which can be considered as “traces” of competences. Often the concepts behind these keywords or phrases are not straightforward but ambiguous. These reasons make it necessary to structure rigorously and carefully craft the ontology, making possible semantic treatments when necessary.

3.1 Methodologies for ontology engineering

The construction of an ontology is a complicated process, requiring the definition of rigorous principles or mechanisms to identify concepts and relationships. Several scientific research efforts have tried to create and formalise ontology engineering methods [7][5]. However, most of these advances are linked to some specific point of view. Importantly, there is still no convergence on the best practices on this issue.

The methodologies proposed to engineer an ontology, structure the transition from non formal knowledge towards a conceptual domain. Then, the formal model of this domain is transformed into an operational ontology. It is commonly accepted to follow these steps of conceptualisation, formalisation, operationalisation in this order. However, while the majority of the methods provide a well structured procedure, most of them still rely on intuitive criteria and on non-formal opinion of experts when modelling the concepts of a knowledge field.

To cope with the ambiguities mentioned earlier, we need to more rigorously conceptualise the notions related to firm competences. Therefore, we selected ARCHONTE methodology [2], which defines precise principles and cognitive mechanisms for each step formalisation of ontology. This method induces a clear definition of the meaning for each concept of the ontology, through a « semantic normalisation » mechanism. ARCHONTE is composed of three main steps illustrated in the following sections: normalisation, formalisation, operationalisation.

3.2 Building an “Ontology for competence traces”

3.2.1. Normalisation

The first normalisation step is to render explicit the concepts of the knowledge domain. These concepts should be expressed in the form of “cognitive constructs”, which have a non-contextual meaning. This allows the arrangement of various constructs into semantically interpretable formulations. [2] proposed to identify such cognitive constructs by applying differentiation mechanisms among concepts. He

suggested applying 4 semantic differentiation principles to clarify the difference between each “cognitive constructs” and its “parent” constructs¹.

Our objective is to engineer an ontology that deals with “competence traces”. The normalisation procedure has been divided into three phases. The first automatically identifies candidate terms for the ontology. A corpus was extracted from mechanical industrial company web sites, with the indexation tool SMART² for this application.

Using this first identification of terms, the second phase uses informal domain expertise, asking several experts to provide an initial version of the potential organisation of the domain concepts (conceptual classes and their relationships). Here the experts’ thoughts have “converged” to two approaches. First, a top-down method offers a generic conceptualisation of the notion of “competence traces”. We show in Figure 2 that the results are a so called “generic ontology”. Second, a bottom up approach which entails a pragmatic notion of competence traces, closely linked to the activity field analysed (mechanical industry). This 2nd approach results in a enumerated part of the final ontology called the “domain ontology”.

Finally, the third normalisation phase reduces ambiguities among the concepts identified and improve the formalisation of relationships among terms and their definitions. The differentiation principles proposed by Bachimont were applied.

3.2.1. Formalisation

The formalisation structures the various levels of the ontology, notably by defining additional properties to the “cognitive constructs” identified previously. The three main types are differentiated: metaphysical, structuring, and parataxic concepts. These 3 types correspond also to distinct generality levels, from the conceptual level of the ontology, to the more pragmatic.

The first level of the “competence traces” ontology is composed of very generic so-called “metaphysical concepts”. For this level, the generic competence model mentioned in section 2 must be created. Two generic classes are defined which correspond to the types of company capabilities of interest: the classes “technological competence traces” and “methodological competence traces”.

The structuring level covers a set of interrelated concepts which should make possible to provide an enumerative and conceptual description of the knowledge domain. Note: here, the “competence traces ontology” CTO is what we called just before the “generic ontology”. CTO structures general conceptual classes of “competence traces”. These concepts have the advantage being independent of the specific activity field addressed: CTO can be re-used for various company activity domains. As illustrated in figure 2, CTO details the 2 classes identified at the metaphysical level. The classes “technological trace capabilities” (as well as methodological competence traces) are specialised by a set of more detailed classes of “competence traces”.

The third, parataxic level of the ontology is constituted by what we called in 3.2.1 the “domain ontology”. The “competence trace” concepts at that level are directly linked to the application field (mechanical industry sector). When applying our

¹ The 4 principles induce to formalize what is different or common among one concept and its paternal / fraternal notions.

² <ftp://ftp.cs.cornell.edu/pub/smart/>

approach to another industrial sector, it is interesting to underline that the structuring level of the ontology remains stable and only the parataxic level has to be modified. Furthermore, the concept classes of this “domain ontology” are characterised by sets of class instances. These instances correspond to specific terms of the domains, used as “competence trace *identifiers*”. As explained in section 4, they are directly used to extract the pertinent information which is composed of the company competence traces. Figure 3 partially illustrates this domain ontology.

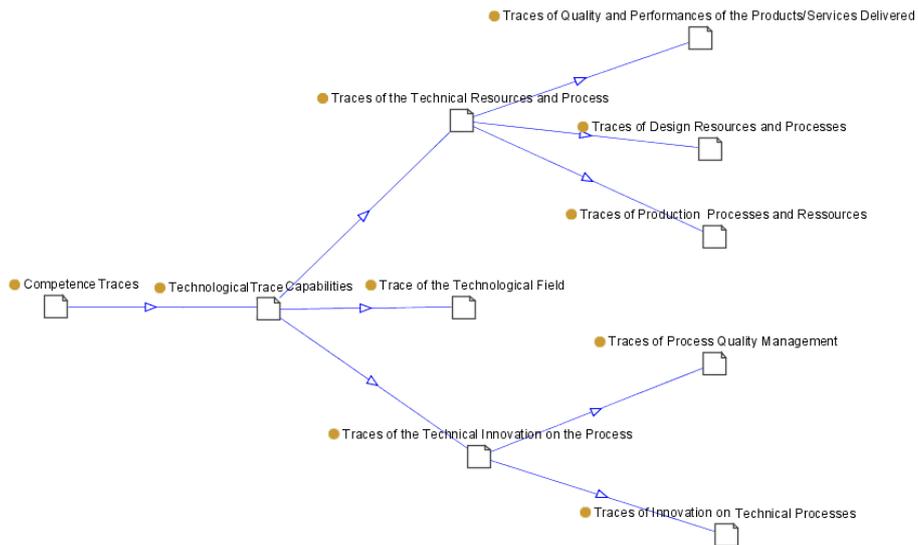


Fig. 2. “Structuring level” of the ontology for “Technological capabilities”

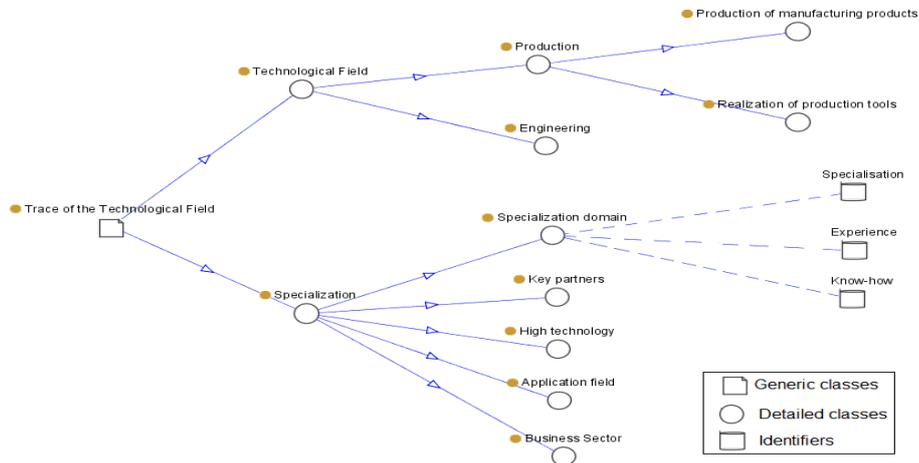


Fig. 3. “Parataxic level” of the ontology (mechanical industry)

Operationalisation will not be described in detail for this paper. This does implement a computer version of the ontology for the application of inference

mechanisms. We have employed the competence trace ontology using *OWL* (*Ontologie Web Language*).

4 Basic Principle of Utilization

As aforementioned, the class instances of the domain ontology are employed as information identifiers. Their role can detect the presence of a given competence trace concept in the web site of a company. A set of identifiers is associated with each conceptual class of the domain ontology to support this validation process. This is called the activation process of the ontology classes. Each company website generates a different activation of the overall ontology, which is interpreted as the competence trace of the company. To avoid ambiguities, the context of use of the identifiers on the web sites has to be considered in the activation process. The performance of the extraction system of competence traces rests on the activation quality of the ontology classes.

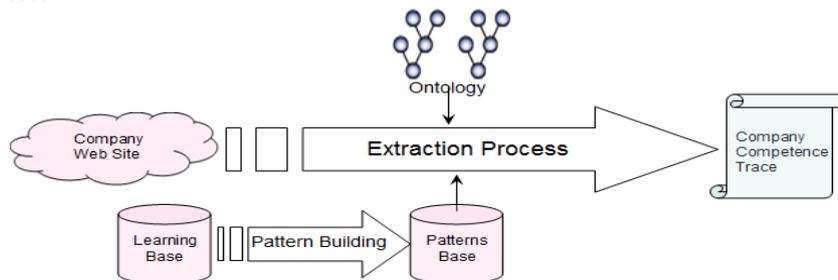


Fig. 4. Extraction process using Ontology and Patterns

To deal with this contextual analysis, a pattern-based approach is currently under development. The approaches by patterns are used in linguistics for associated structural regularities with semantic information [8][6]. Part of the web site corpus, has been used as a “learning base” to identify patterns. The objective is to identify structural regularities linked to the identifiers. Patterns consist in lexico-syntactic expressions which formalise the context of use of these identifiers. In the extraction process of figure 4, patterns are typically used for two objectives:

- To avoid semantic ambiguity when detecting the so-called identifiers.
- To extract additional information on competences, not directly provided by identifiers.

In the learning step which consists in identifying and formalising patterns, a library with roughly 35 elaborated patterns supplemented with 100 unambiguous identifiers (without the necessity of patterns) has been constituted. A specific locating algorithm is under development, which uses these patterns to confirm the presence of identifiers.

Then, depending on the identifiers detected for a given web-site, an activation procedure will confirm or not the activation of each concept class of the competence trace ontology. This will provide the competence traces characterising a company. Afterwards, this information can be further interpreted in a decision aid procedure for establishing VBEs.

5 Discussion and Conclusion

The method followed to engineer a rigorous ontology on competence traces has been explained thoroughly. Future work will build upon this ontology and entail the development of a semantic activation process which requires formalising syntactic patterns. This should be able to measure a similarity among competence traces extracted from various company web sites. This information should then be processed through a decision aid algorithm oriented on network building.

References

1. Berio, G. Harzallah M., Towards an Integrating Architecture for Competence Management, in Special Issue « Competence Management in Industrial Processes », guest editors X.Boucher, E. Bonjour, N.Matta, Computers in Industry, V58, issue 2, February 2007
2. Bachimont B., Isaac A. & Troncy R. (2002). Semantic commitment for designing ontologies: A proposal. In A. Gomez-Pérez & V. Benjamins, Coordinateurs, 13th EKAW Conference, volume LNAI 2473 of LN. in Artificial Intelligence, p. 114–121 : Springer
3. Camarinha-Matos, LM and Afsarmanesh H., *Elements of a base VE infrastructure Computers in industry*, vol. 51, 139-163, 2003.
4. Ermilova E., Afsarmanesh H., *Modeling and management of profiles and competences in VBEs*. Journal of Intelligent Manufacturing 18, 561-586, 2007.
5. Fürst Frédéric. PhD Thesis, Contribution à l'ingénierie des ontologies: une méthode et un outil d'opérationnalisation. Université de Nantes, France, Novembre 2004.
6. M.Finkelstein-Landau and E.Mori. Extracting semantic relationships between terms: Supervised vs. unsupervised methods. In *proc. Of the Int. Workshop on OEGII, Dagstuhl Castle, Germany, pages 71-80, 1999.
7. Grüninger M. et Fox M.S. (1994). *The Role of Competence Questions in Enterprise Engineering*. Paper presented at the IFIP WG 5.7 Workshop on Benchmarking. Theory and Practice, Trondheim, Norway.
8. Hearst M. A. Automatic acquisition of hyponyms from large text corpora. In A. Zampolli, editor, Computational Linguistics (CoLing'1992), pages 539-545, Nantes, France, 1992.
9. Hajlaoui k., Boucher X., Mathieu M., Information Extraction procedure to support the constitution of Virtual Organisations. Research Challenges in Information Science, RCIS'2008, Marrakech, 2008a.
10. Hajlaoui K., Boucher X., Mathieu M., Data Mining To Discover Enterprise Networks, 9 th IFIP Working Conference on Virtual Enterprises (PRO-VE'08) Poznan, POLAND, 8-10 September 2008b.
11. Hodík Jirí, Jirí Vokřínek, Jirí Bíba, Petr Becvár: Competences and Profiles Management for Virtual Organizations Creation. CEEMAS 2007: 93-102.
12. Pepiot G., PhD Thesis, Modélisation des Entreprises sur la base des compétences, EPFL, 2005.
13. Plisson, J. Ljubic P, Mozetic I, Lavrac N. (2007). An ontology for Virtual Organisation Breeding Environments. *To appear in IEEE Trans. On Systems, Man, and Cybernetics*.
14. Yussopova Y. & Probst A.R., *Business concepts ontology for an enterprise performance and competences management*, in Special Issue « Competence Management in Industrial Processes », guest editors X.Boucher, E. Bonjour, N.Matta, Computers in Industry, V58, issue 2, February 2007.
15. Wernerfelt B., A Resource-Based View of the Firm, *Strategic Management Journal* 5, 171-180, 1984.