Ontology Network for Social Network Analysis in a Knowledge Management Context

Bárbaro Gonzalez[1], Omar Chiotti[1] and Mariel Ale[2]

[1]INGAR-CONICET
Avellaneda 3657, Santa Fe, Argentina
barbaro@santafe-conicet.gov.ar,
chiotti@santafe-conicet.gov.ar,
[2] CIDISI - UTN
Lavaisse 610, Santa Fe, Argentina
male@utn.frsf.edu.ar

Abstract. Organizational knowledge is one of the most valuable assets that companies own today. For several decades organizations have been developing strategies to manage knowledge with particular emphasis on tacit knowledge discovery. The particular dynamic that presents the evolution and transfer of tacit knowledge is closely tied to the relations between people. For this reason, Social Network Analysis (SNA) can be a powerful tool to support a Knowledge Management (KM) initiative. Despite usefulness recognition of SNA techniques within KM processes, there is still remains the initial problem of data collection and representation (problem shared by both initiatives). The aim of this paper is to analyze an ontology network usefulness to obtain the necessary knowledge structure to feed the SNA-KM integration architecture proposed.

Keywords: Ontology Network; Social Network Analysis; Knowledge Management

1 INTRODUCTION

The particular dynamic that presents the evolution and transfer of tacit knowledge is closely tied to the relations between people within the organization who are the main containers of this type of knowledge. For this reason, Social Network Analysis (SNA) can be a powerful tool to support a Knowledge Management (KM) initiative [1].Any KM initiative should be nurtured not only of relations between individuals but also between individuals and Knowledge Objects (KO) that are vital for business development within the organization.

In recent decades, SNA has become a formally established research method for social science with dedicated journals (Social Networks, Journal of Social Structure and Connections), textbooks and handbooks [2], specific software (Ucinet) and an association (the International Network for Social Network Analysis). All of this has allowed SNA to spread to other disciplinary fields generating several lines of research.

The first concern before any social network analysis is the collection of primary data on which the study will be conducted. Traditional SNA methods based on interviews and surveys have proved useful for obtaining a basis for understanding information communication and transfer in social networks. However, most studies using these techniques have been limited to relatively small data sets mainly due to difficulties in network members' access, the time and effort required for participants to complete the questionnaires and ethical, analysis and interpretation issues.

The growth of online interactions within the business world and the recording of these interactions opened a new data source for SNA techniques application. Using this data source has many KM related advantages for the organization. On the one hand, it reflects the dynamics of organizational knowledge evolution and, an analysis of these interactions allow inferring topics of interest (knowledge objects) for the organization and who are those who know about these issues (referrals). On the other hand, its automatic collection is aligned with one of the basic prerequisites for success of any KM initiative related with not to impose a work overload to workers.

The aim is to analyze ontologies usefulness to obtain the necessary knowledge structure to feed the SNA-KM integration architecture proposed. This paper focuses on the development of the knowledge layer of this architecture. To this end, the paper is organized as follows. Section 2 presents a review of previous works related to KM and SNA integration. Section 3 outlines the SNA-KM integration architecture proposed. The ontology based knowledge layer is presented in Section 4. Finally, Section 5 presents conclusions and future challenges in the area.

2 RELATED WORK

Analysis and discussion of network structures and its influence on management was strongly influenced by Drucker [3], Savage [4] and later by Kanter [5] which emphasized the importance of networks in knowledge management and distribution. According to these authors, organizations that develop and promote both internal and external networks are in a better position when it comes to managing their knowledge.

However, despite the recognition of networks as the ideal means for organizational knowledge creation and distribution, nor a systematic development of methods for networks and knowledge communities recognition, neither the analysis of their structure and evolution that allow a practical use of them is observed. It is at this point that the SNA methods may become a useful tool for KM.

SNA has made important contributions to a variety of fields including epidemiology, anthropology, social psychology, etc. However, the application of SNA techniques to KM or knowledge modeling itself is relatively new.

The link between KM and SNA techniques was traditionally related to recommender systems [6]. Such systems seek to predict the 'rating' or 'preference' that a user would give to an item (such as music, books, or movies) or social element (e.g. people or groups) they had not yet considered, using a model built from the characteristics of an item (content-based approaches) or users' social environment (collaborative filtering approaches) [7].

The idea of the usefulness of SNA in activities related to knowledge is based on the notion that social networking is a key factor in understanding knowledge creation processes. Hildreth and Kimble [8] suggest that knowledge creation and social networks are closely related and that this relationship has a positive connotation. These networks also represent relationships between members and the availability and exchange of knowledge resources in the network [9].

Other authors who delved into organizational dynamics indicated that knowledge distribution requires social processes and interactions usually due to the tacit nature of knowledge [10]. In this context, applying SNA techniques seems natural. Nonaka and Takeuchi [11] also argued that some level of co-presence, social affinity, and socialization are required to enable effective transfer of knowledge that is difficult to codify. Knowledge creation is a collaborative process by which domain members interact, develop, and exchange new knowledge while shaping the formal and informal networks of a particular domain [12]. In fact, social networks facilitate knowledge creation process because they define connectivity of members, which in turn directly affects the conditions of intellectual collaboration and exchange processes between members. Studying social networks, thus, has become a major organizational focus on developing partnerships in communities where the network is constituted by the key processes in knowledge creation and distribution [13]. These are key processes to any initiative of organizational knowledge management.

The effectiveness of an organization and its ability to accomplish its full operational potential largely relies on the strength of the relationships between its individuals and the presence of multiple knowledge flows. However, little analysis has been done on other relationships that are critical when managing knowledge in an organization such as those between people a KO and people and tasks.

Despite usefulness recognition of SNA techniques within KM processes, there is still remains the initial problem of data collection (problem shared by both initiatives). Therefore, it is not surprising that there is growing interest in automatically recovering and analyzing individuals' online behavior using Web and data mining techniques [14].

3 KM AND SNA INTEGRATION

In the KM area, SNA techniques could give support to three main areas: the discovery of an informal structure that coexists with the formal structure within the organization (this occurs even in larger rigid and bureaucratic companies), the manifestation of knowledge resources (individuals or objects) that are critical or central to the organization, and the facilitation of location and access to these resources.

Networks formation within an organization has important implications for all aspects of organizational life. Numerous network theoretical models and empirical studies have examined how the network structure affects the results of a variety of tasks [15].In this context, SNA is shown as a promising approach to help organizations manage a number of classic situations including leadership and task force selection, informal structure discovery and knowledge resources manifestation among others

[16].Ale and Galli proposed different perspectives relating people, knowledge objects, and tasks in an integration SNA-KM architecture (Figure 1) [1]. These perspectives were defined in relation to three layers that must be taken into account when implementing a KM initiative: social layer (containing individuals within the organization), knowledge layer (containing all those knowledge objects valuable to the organization along with its classification structure - an ontology in our case), and the business process layer (containing the set of business process models of the organization).

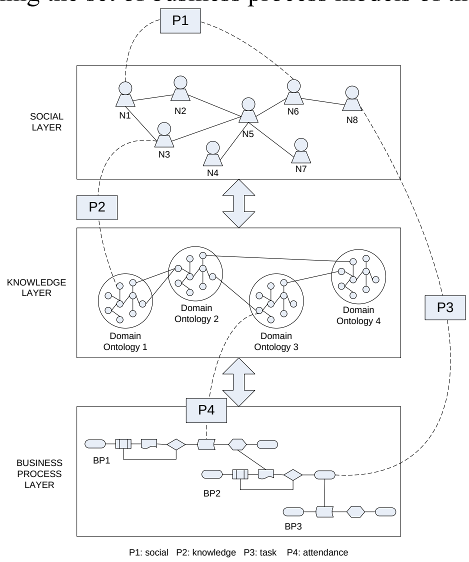

Fig.1. Integration perspectives

The commonly used perspective, the social one, aims to represent relationships between individuals within the organization (P1). The knowledge perspective is intended to determine who knows what within the organization and therefore it relates knowledge objects with people who know them. This perspective is defined between the social layer and the knowledge layer (P2). The knowledge layer contains organizational knowledge distributed on different domains that correspond to different business units or departments in the organization. The task perspective relates people and tasks or events that should be included as part of their daily work (P3). The attendance perspective aims to answer the question of what knowledge objects are required to perform a given task or pursue a particular event (P4). These perspectives are related with the business process layer. This layer contains the business processes defined within the organization. An analysis of relationships across these perspectives may point to the need for change in the organizational structure or to the granting of new roles to solve identified bottlenecks. Figure 1 shows the four perspectives identified through the layers defined in the organization [1]. Once integration perspectives between SNA and KM are defined is necessary to analyze the possibility of collecting the necessary data to feed the proposed architecture and thus take advantage of such integration.

Due to the advancement of data mining techniques the main components of the social layer (nodes and relations) can be extracted and characterized [17]. A technique commonly used for node detection involves the discovery of names and references to persons within the text. There are two main approaches to this task, the first involves the search for names that are in a dictionary [18] the second applies linguistic rules or patterns to the content and sentence structure to identify potential names. These patterns and linguistic rules are usually built based on characteristic attributes of words such as frequency, context and position in the text [19]. For details on the recognition of named entities see Nadeau and Sekine [20].
Knowledge involved in organizational domains (knowledge layer) can be classified into three groups: knowledge commonly used throughout the organization, knowledge specifically used by a domain and knowledge shared between domains. This knowledge classification can be properly represented by an ontology network, in which common concepts are modeled by a general ontology, domain specific concepts are modeled by domain ontologies and concepts shared between domains are modeled by relationships between concepts of domain ontologies.

4 KNOWLEDGE LAYER DESIGN

As it was said in Section 1, this work focuses on the development of the knowledge layer. To carry out the design and development of this layer the NeOn methodology has been applied, using NeOn Toolkit1. This particular methodology has been chosen due to NeOn [21] has been successfully used to build ontology networks for different domains and by people with diverse background, for example, and just to name a few, in e-employment [22], in education [23] in tourism [24] and in mobile environments [25]. NeOn can be adapted to user needs, and includes new processes and activities involved in developing ontology networks. Following this methodology, the competency questions technique [26] to elicit ontology requirements was used.
Enron Corpus [27] is used as information source to discover the network of domains ontologies. The Enron Corpus contains 96,107 messages from the "Sent Mail" directories of all the users in the corpus. Divided across 45 files, the Enron Corpus contains 2,205,910 lines and 13,810,266 words. This corpus was selected due to it is the single corpus of "real" emails publicly available suitable for this study.

4.1 Ontology Network Scope and Formality Level

Knowledge management ensures the development and application of all types of relevant knowledge in a company in order to improve their ability to solve problems and contribute to the sustainability of their competitive advantages. In this context, it is crucial to identify those who are a source of valuable information and support the transformation of organizational knowledge to some structured form that can be processed. Consequently, the purpose of the ontology network proposed is to model

[1] http://neon-toolkit.org/wiki/Main_Page.html

semantic information concerning the knowledge objects of Enron employees from the data obtained in the Enron Corpus. In terms of scope, services provided by Enron are grouped into the following domains: Energy services, Broadband services, Risk Management services and General specification. The last domain involves those concepts that are common for the three services type.

To meet the requirements of the platform proposed the ontology network is implemented in OWL 2. This language provides a good balance between expressivity and computational completeness.

4.2 Requirements and Intended users/uses identified

The analysis of the motivating scenarios allowed us to identify as intended users of the ontology: managers and employees that create/use knowledge objects. The analysis of the motivating scenarios described in [21], allowed us to identify as the main intended uses of the ontology: representation, search and retrieval of knowledge objects.

As non-functional requirements we can mention that the ontology has to be developed in English.

Due to the lack of domain experts, competency questions were elicited from Enron Corpus. Forty competency questions grouped into four groups were defined:

Energy services Group: covers all the topics and concepts of one of the business units within Enron, whose purpose was to supply gas, electricity and related management services directly to businesses and homes.

Broadband services Group: covers all the topics and concepts of three business units within Enron: intermediation, fiber optic network and content services. Acquisitions business was based on the e-commerce platform Enron Online.

Risk management services Group: covers all the topics and concepts of a business unit that was used by both Enron and their customers. The purpose of this unit was managing risks to provide security and integrity of the buying and selling of products. Financial accounting and commercial prepaid were part of the service they provided.

General Group: Common themes and concepts shared with other groups.

Some of the competency questions of General Domain can be seen in Table 1.

Competency Questions – General Domain
1. What are the research projects?
2. Who are the employees of Enron that participate in research projects?
3. How many employees does Enron have?
4. What are the main services provided by Enron?
5. What are the areas of knowledge generated by employees in the services?
6. How many customers have by services?
7. What are the contracts that Enron have?
8. How much annual revenue generated by the Company's services?
9. What are the Enron departments?
10. What are the most important departments by services?

11. What are the most qualified employees by departments?
12. What are the filed reports?
13. What are the Enron power plants?
14. What countries have the power plants?

Table 1. Examples of competency questions from General Domain

4.3 Terminology extraction

Competency questions were answered by performing a text mining of the Enron Corpus by using Automap[2]..Main terms were extracted from competency questions and their answers, counting frequency of their occurrence. Extracted terms were formally represented in the ontology network by means of concepts, attributes and relations.

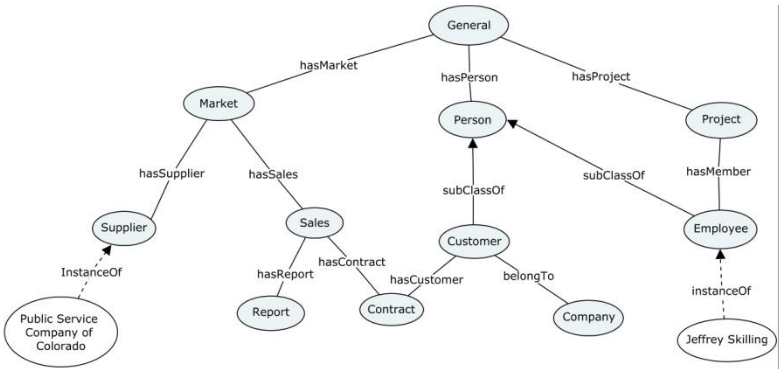

Fig.2. Ontology Domain – General

In Figure 2 main concepts of domain General can be seen. Term *General* represents the ontology concept "Thing". It is related to *Market*, *Person* and *Project* concepts through *hasMarket*, *hasPerson* and *hasProject* relationships respectively. Relationship *hasMember* relates *Employee* and *Project* concepts denoting that employees are part of company projects. Concept *Employee* is a subclass of concept *Person* and inherits its properties. Relationships *hasSupplier* and *hasSales* relate the *Market* concept with *Supplier* and *Sales* concepts. Relationships *hasContract* and *hasReport* relate *Sales* concept with *Contract* and *Report* concepts respectively. These relationships denote that sales are made through contracts and informed through reports.
Relationship *hasCustomer* relates *Contract* concept to *Customer*. The *Customer* concept is related to *Person* and *Company* concepts through *subClassOf* and *belongTo* relationships respectively. Relationship *subClassOf* denotes that customer is a subclass of person and inherits its properties and relationship *belongTo* denotes that a customer belongs to a company.

[2]http://www.casos.cs.cmu.edu/projects/automap/

4.4 Ontology Network

The same procedure applied to obtain the General ontology for domain General specification was followed to obtain the ontology for services domains: Energy services ontology for domain Energy services, Broadband services ontology for domain Broadband services, and Risk Management ontology for domain Risk Management services (Top of Figure 3). The ontology network is conformed through relationships between the General ontology and the ontologies of three services domains provided by Enron.

Relationship usesSymbolsOf [28] (Figure 3) denotes that properties of an ontology concept involve instances of a concept from other ontology. So, even though they are separated ontologies, an ontology depends on another due to the first has properties involving instances of the second. Relationship semanticallyIncludedIn[29] relates common concepts between two ontologies. So, Energy services ontology, Broadband services ontology and Risk Management services ontology are related with the General ontology through a semanticallyIncludedIn relationship (Top of Figure 3).

For example, relationship semanticallyIncludedIn between General ontology and the Risk Management ontology allows relating (through relationship hasInvestment) Customer concept modeled by the General ontology with Investment concept modeled by the Risk Management ontology (bottom of Figure 3). This relationship represents that the investment funds managed by the Risk Management services are owned by the customer. Relationship semanticallyIncludedIn between General ontology and Broadband services ontology allows relating (through relationship employeeOf) Employee concept modeled by General ontology with Broadband services concept modeled by Broadband services ontology (bottom of Figure 3). This relationship represents that the broadband service has employees. For the same purpose, Employee concept is also related through employeeOf relationship to Energy service and Risk management, which are the main concepts of Energy services ontology and Risk Management ontology respectively.

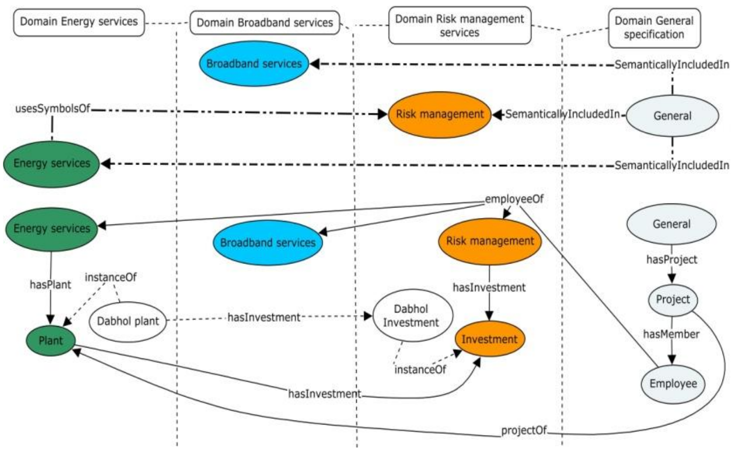

Fig. 3. Ontology Network

Relationship s*emanticallyIncludedIn* between General ontology and Energy services ontology allows relating (through relationship *projectOf*) *Project* concept modeled by General ontology with *Plant* concept modeled by Energy services ontology (bottom of Figure 3). This relationship represents those projects that involve Enron production plants. Relationship *usesSymbolsOf* between Risk Management ontology and Energy services ontology allows relating (through relationship *hasInvestment*) *Plant* concept modeled by Energy services ontology with *Investment* concept modeled by Risk Management ontology. This relationship represents that the instances of *Plant* concept have *Investment* among their properties.

5 CONCLUSIONS AND FUTURE WORK

In this paper, the topic of ontology network design to feed the proposed methods for SNA-KM integration architecture was addressed. The main focus was to develop an knowledge representation model to support the knowledge layer defined in the integration architecture of enterprise knowledge management and social network analysis techniques. The ontology network proposed constitutes a representation layer of organizational knowledge that provides a homogeneous view of organizational knowledge objects that are naturally heterogeneous.

The modeling of knowledge objects enables to define some perspectives of the proposed architecture. Specifically, those perspectives that relate knowledge objects with people and knowledge objects with tasks.

As future work, there is the testing of techniques for knowledge objects classification using the ontology network defined with the appropriate search and retrieval strategies of these objects and the modeling of the other two architecture layers.

REFERENCES

[1] M. Ale and M.R. Galli, "Integration Perspectives of Organizational Knowledge Management and Social Network Analysis", *Proceedings of XXXIX Latin American Computing Conference*, CLEI 2013, pp. 414-421.

[2] J. Scott and P. Carrington, Eds., *The SAGE Handbook of Social Network Analysis [Hardcover]*. SAGE Publications Ltd, 2011, p. 640.

[3] P. Drucker, "What business can learn from nonprofits," *Harvard business review*, vol. 67, no. 4, pp. 88–93, 1989.

[4] C. M. Savage, *5th generation management: integrating enterprises through human networking*. Newton, MA, USA: Digital Press, 1990, p. 267.

[5] R. M. Kanter, *Evolve!: Succeeding in the Digital Culture of Tomorrow*. Harvard Business Press, 2001, p. 352.

[6] Y.-M. Li, T.-F. Liao, and C.-Y. Lai, "A social recommender mechanism for improving knowledge sharing in online forums," *Information Processing & Management*, vol. 48, no. 5, pp. 978–994, Sep. 2012.

[7] F. Ricci, L. Rokach, and B. Shapira, Eds., *Introduction to recommender systems handbook*. Springer US, 2011, pp. 1–35.

[8] P. Hildreth and C. Kimble, *Knowledge networks: Innovation through communities of practice*. IGI Global, 2004, p. 354.

[9] C. Haythornthwaite, “Social network analysis: An approach and technique for the study of information exchange,” *Library & Information Science Research*, vol. 18, no. 4, pp. 323–342, Sep. 1996.

[10] R. Gulati, “Does familiarity breed trust? The implications of repeated ties for contractual choice in alliances,” *Academy of management journal*, vol. 38, no. 1, pp. 85–112, 1995.

[11] I. Nonaka and H. Takeuchi, *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press, 1995, p. 304.

[12] M. A. McFadyen, M. Semadeni, and A. A. Cannella, “Value of strong ties to disconnected others: Examining knowledge creation in biomedicine,” *Organization science*, vol. 20, no. 3, pp. 552–564, 2009.

[13] D. Klimkeit, “Networks , Forms of Exchange and Motivations. Insights from Social Anthropology for the Issue of Knowledge Sharing,” in *Professional Knowledge Management*, 2005, pp. 392–397.

[14] A. Gruzd and C. Haythornthwaite, “Networking online: cybercommunities,” in *The SAGE Handbook of Social Network Analysis*, J. Scott and P. J. Carrington, Eds. SAGE Publications Ltd, 2011, pp. 167–179.

[15] D. Enemark, M. McCubbins, and N. Weller, “Knowledge and networks: An experimental test of how network knowledge affects coordination,” *Soc. Networks*, 2014.

[16] R. Cross, A. Parker, and S. Borgatti, “A bird’s-eye view: Using social network analysis to improve knowledge creation and sharing,” 2002. [Online]. Available: http://www.workinfo.org/index.php/articles/item/709-a-bird-s-eye-view-using-social-network-analysis-to-improve-knowledge-creation-and-sharing. [Accessed: 25-Apr-2014].

[17] A. Gruzd and C. Haythornthwaite, “Networking online: cybercommunities,” in *Handbook of Social Network Analysis.*, J. Scott and P. Carrington, Eds. SAGE Publications Ltd, 2011, pp. 168–179.

[18] M. Harada, S. Sato, and K. Kazama, “Finding authoritative people from the web,” *Proc. 4th ACM/IEEE-CS Jt. Conf. Digit. Libr.*, pp. 306–313, 2004.

[19] D. Nadeau, P. Turney, and S. Matwin, “Unsupervised named-entity recognition: Generating gazetteers and resolving ambiguity,” in *Lecture Notes in Computer Science: Advances in Artificial Intelligence*, J. G. Carbonell and J. Siekmann, Eds. SAGE Publications, Inc; 2nd edition, 2006, pp. 266–277.

[20] D. Nadeau and S. Sekine, “A survey of named entity recognition and classification,” *Lingvisticae Investig.*, vol. 1, no. 30, pp. 3–26, 2007.

[21] M. C. Suárez-Figueroa, A. Gómez-Pérez, and M. Fernández-López, "The NeOn Methodology for Ontology Engineering", *Book Chapter in Ontology Engineering in a Networked World*, 2012, Publisher: Springer Berlin Heidelberg, pp. 9-34.

[22] Villazón-Terrazas B, Ramírez J, Suárez-Figueroa MC, Gómez-Pérez A (2011) A network of ontology networks for building e-employment advanced systems. *Expert Syst Appl* 38(11):13612–13624.

[23] Clemente J, Ramírez A, de Antonio A (2011) A proposal for student modeling based on ontologies and diagnosis rules. *Expert Syst* Appl 38(7):8066–8078.

[24] Lamsfus C, Alzua-Sorzabal A, Martin D, Salvador Z, Usandizaga A (2009) Human-centric ontology-based context modelling in tourism. In: *Proceedings of KEOD 2009 Proceedings of the International Conference onKnowledge Engineering and Ontology Development*, Funchal - Madeira, Portugal, October 6–8, 2009. INSTICC Press 2009, ISBN 978-989-674-012-2, pp 424–434.

[25] Poveda-Villalón M, Suárez-Figueroa MC, García-Castro R, Gómez-Pérez A (2010) A context ontology for mobile environments. In: *Workshop on Context, Information and Ontologies (CIAO 2010)* co-located with EKAW 2010, Lisbon.

[26] M. Gruninger, M. S. Fox. The role of competency question in enterprise engineering. In Proceedings of the IFIP WG5*.7 Workshop on Benchmarking - Theory and Practice*, Trondheim, Norway, 1994.

[27] The Enron Sent Corpus v1.0. Compiled by Will Styler at the University of Colorado. Available at http://www.cs.cmu.edu/~enron/.

[28] Dıaz, A., Motz, R., & Rohrer, E. (2011). Making ontology relationships explicit in a ontology network. *In V Alberto Mendelzon International Workshop on Foundations of Data Management* (AWM 2011), Santiago, Chile.

[29] (C. Allocca, M. D’Aquin, and E. Motta. DOOR - Towards a Formalization of Ontology Relations. In Jan L. G. Dietz, editor, KEOD, pages 13–20. INSTICC Press, 2009.).