A collaborative methodology for tacit knowledge management: Application to scientific research

Emna Mezghani [a,b,c,d,*], Ernesto Exposito [a,c], Khalil Drira [a,b]

[a] CNRS, LAAS, 7 avenue du colonel Roche, F-31400 Toulouse, France
[b] Univ de Toulouse, LAAS, F-31400 Toulouse, France
[c] Univ de Toulouse, LAAS, INSA, F-31400 Toulouse, France
[d] Luxembourg Institute of Science and Technology, 5, Avenue des Hauts-Fourneaux, L-4362, Esch/Alzette, Luxembourg

HIGHLIGHTS

- Proposing a methodology for formalizing tacit knowledge to cover social barriers.
- Applying the proposed methodology to the scientific research management.
- Defining the Network and Communication Research Domain Ontology (NCRDO).
- Implementing and evaluating a collaborative semantic web platform based on NCRDO.

GRAPHICAL ABSTRACT
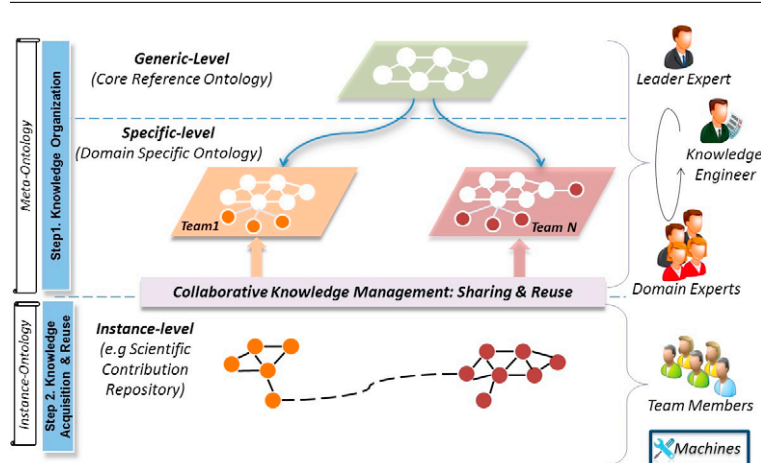
ARTICLE INFO

ABSTRACT

Tacit knowledge, which refers to the know-how, is critical to understand and reuse since it is located in the human heads. It represents the foremost element for human and team evaluation. Seeking for tacit knowledge is achieved only by communicating with the concerned persons, which makes losing it axiomatic if people leave their work without documenting their know-how. Thus, providing a collaborative environment based on a common conceptualization of the domain to formalize the experts' knowledge and to share their outcomes is required. However, some barriers pertaining to cultural and social factors such as personality traits impede capturing the conceptual model. To cope with these issues, we have proposed a generic two-step methodology that copes with human barriers when capturing the domain experts' tacit knowledge, their skills, and seeds terms in order to converge to a common knowledge representation. Considering the scientific research management as a use case, we followed the proposed methodology to formalize our scientific research knowledge in the context of network and communication research field. Based on the generated ontology, we have developed a semantic web platform that allows collaboratively annotating experts' knowledge in a computer interpretable format

* Corresponding author at: CNRS, LAAS, 7 avenue du colonel Roche, F-31400 Toulouse, France. Tel.: +33 561336921.
  E-mail addresses: emna.mezghani@laas.fr (E. Mezghani), ernesto.exposito@laas.fr (E. Exposito), khalil.drira@laas.fr (K. Drira).

that can be shared and reused by human and machines. Our evaluation is based on end users' quality of experience and feedbacks.

1. Introduction

Research collaboration plays an important role in evaluating the research team outcomes [1,2]. Successful collaborations take place if research team members and/or external partners can smoothly communicate based on a sharable knowledge that reflects the team skills [3]. Currently, scientific research team knowledge is embedded in unstructured documents including published articles, internal team reports, figures, tables and videos representing demonstrations. Thus, information is everywhere, heterogeneous and poorly organized which hamper knowledge exchange and sharing, and make information seeking is time-consuming.

Advances in Knowledge Management (KM) based on Information Technology (IT) provide important means to increase productivity and achieve the team effectiveness [4] since it provides methods and tools that capture, understand, share, and facilitate knowledge access and reuse by team members to create value [5–7]. However, IT support cannot cover more than 10%–30% of KM [8]. More importantly, domain experts are the main knowledge source. Face-to-face interactions such as interviews, brainstorming, meetings, etc. are the keys for elucidating, capturing and sharing experts' tacit knowledge [9]. It is worth mentioning that the success of KM, especially which relies on the knowledge of experts, strongly depends on the acceptance of people involved in this process. At this end, human collaboration and participation is required to underpin an efficient knowledge transfer and sharing. Nevertheless, some cultural and social factors may hinder the progress of this process, especially if members may feel and think that knowledge sharing depletes the time and the efforts that can be invested in other activities more beneficial for themselves [10]. Moreover, each member has his/her own personality traits and characteristics, which makes managing a group of person complex, in particular if conflicts appear.

Consequently, to handle human barriers that may hamper extracting and sharing the experts' tacit knowledge, we propose a generic two-step methodology that describes processes and actors involved in the conceptualization and knowledge creation based on a multi-level approach. The objective of proposing a multi-level approach is to identify as a first stage a generic domain level that can be reused by any community/group of users interested to this domain, then extending it to support the user-requirements. We applied the proposed methodology to formalize and classify the scientific research activities within our SARA research team.[1]

The rest of the paper is organized as follow. Section 2 draws our proposed generic two-step methodology for experts' tacit knowledge conceptualization. Section 3 provides the results of applying the proposed methodology to formalize the tacit knowledge of our team experts. Section 4 presents our preliminary evaluation of the efficiency of the proposed methodology. Section 5 details existing researches dealing with collaborative scientific knowledge management and describes existing methodologies for building ontologies. Finally, Section 6 summaries our contribution and identifies our future research works.

---

[1] Services and Architectures for Advanced Networks- SARA: https://www.laas.fr/public/fr/sara.

2. A generic two-step methodology for building ontology

Our methodology follows a top-down approach [11] which is mainly composed on two steps as shown in Fig. 1: the "*Knowledge Organization*" that corresponds to conceptualizing and representing the knowledge in an appropriate format; and the "*Knowledge Acquisition and Reuse*" that allows the users collaboratively producing and consuming the knowledge. The "*Knowledge Organization*" represents the fundamental step. First, it identifies a core ontological model named *Core Reference Ontology (CRO)* describing the generic concepts and relations according to the formalized requirements. Then, it specializes these concepts to reflect the domain experts' skills and knowledge based on their collaboration, and generates the *Domain Specific Ontology (DSO)*. The main contribution of this methodology is to avoid conflicts and human barriers related to the personality characteristics when communicating with experts.

Based on an incremental approach, requirements are captured and refined with the collaboration of the leader expert. Thus, the key concepts of *CRO*, inference and inconsistency rules are identified (1, 2). After preparing the list of concepts and relations, the knowledge engineer identifies existing upper ontologies that can be reused and easily integrated with the *CRO* (3). To personalize the *CRO* in a specific domain, the domain experts are involved to enrich the *CRO* with their own vision and tacit knowledge.

It is truism that each person has his/her specific character and manners to communicate and express his/her knowledge. Studies in psychology have shown that knowledge sharing behavior among individuals is influenced by personality traits [12]. According to Pervin [13], personality refers "*to an individual's unique and stable pattern of thinking, feeling, acting and reacting to his or her social environment*". These patterns, which explain why some individuals are motivated to share knowledge while others are not, can be classified according to the big five taxonomy [14]: *extraversion* that refers to energetic people who prioritize harmony in their social relationships; *agreeableness* that concerns friendly and cooperative people; *conscientiousness* that refers to reliable and responsible people; *openness* to experience that encompasses imaginative and flexible people who are interested in learning new experiences; and *neuroticism* that concerns people who may rapidly experience unstable emotions such as nervosity and anxiety. Accordingly, to avoid barriers related to social and psychological dimensions, our methodology adopts an incremental approach in which the identification of the knowledge structure is decomposed into sub-steps. First, experts are split into groups (4). Each group includes experts working on the same research area. Then, for each group, individual meetings with each expert are organized to discuss with him/her, to formalize his/her vision and to acquire as much as possible knowledge about his/her work (5, 6). By this way, problems such as being influenced by colleagues high in extraversion, and avoiding conflicts with colleagues high in neuroticism are covered. Furthermore, according to the expert character, different methods and questions are adopted to extract his/her knowledge. For example, the knowledge engineer reviews some relevant publications of the expert and extracts a set of keywords that help animating the discussion in order to deal with problems pertaining to people less in agreeableness or less in extraversion.

At this stage, the *CRO* is used when discussing with experts to converge to a unified model. By repeating the processes (5, 6), different models are generated.

Afterward, for each group, the different versions are aggregated (7) to provide a unified ontology schema which is approved and validated by the experts of this group (8, 8′). The final versions of all the groups are integrated in order to generate the *Domain Specific Ontology (DSO)* that reflects the team directions and interests, and formalizes its experts' visions (9).

During the design process, we used UML diagrams to model and visualize the ontology classes and relationships since it is accepted and understood by experts. The resulted ontology described in UML will be translated (10) into a formal ontology language such as OWL, RDF and RDFS to drive the knowledge acquisition process and enable reasoning for smart knowledge management. Our platform includes *inference rules* that automatically create new knowledge, and a *Consistency Engine* that detects knowledge inconsistency and incompleteness (e.g. non-compliance of cardinalities) based on domain rules. In particular, our platform relies on semantic web technology and provides a user friendly interface that enables a collaborative authoring and annotation among the team members. It is a kind of virtual knowledge mediator that interconnects team knowledge in a way that can be smoothly reused.

3. Use case application: scientific research management

We followed our proposed methodology to formalize the scientific research domain. As a first stage, we identified the *Scientific Research Activities Core Ontology (SRACO)* that describes the generic concepts related to the scientific research management independent from the team interests. At a second stage, we specialized the *SRACO* to formalize the SARA research team skills and expertise in the context Network and Communication domain. The result of applying this methodology is the *Network and Communication Research Domain Ontology (NCRDO)* depicted in Fig. 2. During the conceptualization phase, we have used the *Astah Community* tool[2] to represent the ontology classes and relationships as UML class diagram.

First, by collaborating with the research team leader, we identified the main administrative and high level scientific concepts such as the "Scientific Contribution", "Approach", "Method", "Technique", etc. Then, we identified possible existing upper ontologies that can be reused such as FOAF[3] ontology that describes the organization structure (e.g. *foaf:Person* and *foaf:Organization*), and the Event Ontology[4] that describes the *event:Event* class. Based on an incremental approach, we aggregated the different identified concepts to generate *SRACO*. The completed version of the *SRACO* implemented in OWL can be downloaded from [15]. The next step is to personalize this ontology to capture the team experts' tacit knowledge and formalize their skills. Herein, our methodology intervenes to solve social issues that may appear with a group of persons. We split the experts into two groups: the "*Network performance*" and the "*Distributed services and applications*". Then, we individually met each expert and applied an incremental process to identify their seeds terms. Different visions and interpretations have been discussed and reformulated in order to converge to the specialized ontology (*NCRDO*) that details the scientific contributions criteria of the SARA team. Fortunately, our experts are with high conscientiousness and openness expressed through their reliability and their ability to help us aligning their knowledge with their colleague's experience in order to generate a unified ontology.

Fig. 2 presents an excerpt of the main classes, relationships and cardinalities of the *NCRDO*. The specialized part is represented as sub-classes of the *SRACO* classes, which are represented in red color in Fig. 2, except the "Research Directions" and the "Applicative Domain" are instances. For clarity reasons, we present only the "Applicative Domain" with the ⟨⟨*instance*⟩⟩ stereotype. A detailed OWL version of the *NCRDO* representing the SARA team interests can be downloaded from [16].

It is noteworthy that our proposed methodology is being used also in the medical domain to formalize the experts' medical knowledge and encode the disease management strategies, based on the collaboration of medical experts, in order to adapt the patient treatment plan [17].

4. Evaluation

To evaluate the efficiency of the proposed methodology, we developed a proof of concept implementing the *Network and Communication Research Domain Ontology (NCRDO)*. We are interested in providing a scalable platform able to support the development of the personalized ontology and the collaborative knowledge annotation to provide the right seeds terms. Consequently, we selected Semantic MediaWiki (SMW) [18] as a semantic web technology since it unleashes the power of wikis for collaborative knowledge management, and ontologies for providing a common understanding of the domain. Moreover, SMW belongs to the "wiki for semantics" category of semantic wiki engines that support collaborative ontology development [19]. It offers mechanisms that map annotations embedded in wiki pages into a formal description using the OWL DL ontology language [20]. Thus, knowledge is collaboratively populated and can be seamlessly used by human and machines.

As a first-line evaluation, we proposed three criteria that reflect the users' Quality of Experience (QoE) presented in Fig. 3. The first criterion is the completeness of the *Domain Specific Ontology*, which refers in this paper to the *NCRDO*. It allows evaluating the efficiency of our proposed methodology to extract as much as possible knowledge from the experts. Thus, the other research team members (PhD, PostDoc, Visitor, Trainee, etc.) can easily annotate their scientific contributions. The second criterion is more related to the openness of the platform to support the evolution of the *NCRDO* from both the meta-level and instance-level (e.g. adding new seeds terms with the respect of the *SRACO* schema). Finally, the third criterion refers to evaluating the human–machine interaction when (1) acquiring the knowledge and, (2) automatically generating new knowledge based on the semantic search mechanisms. As advocated in Fig. 3, the evaluation provides a good sight about the proposed methodology in providing an ontology covering the main keywords of the team interests and directions.

5. Related work

Ontology has been widely used in different domains to ensure collaboration and provide common structure to knowledge for an easy share and reuse. Pirró et al. [21] described a framework implementing distributed ontology-based knowledge management systems within organization. The authors proposed a multi-level approach that identifies an upper ontology and a set of Core Ontologies Knowledge Entities describing the basic knowledge background of an organization, then deepens it to a particular domain in which an individual knowledge worker is interested. In the context of e-learning, Shih et al. [22] proposed an ontology-based approach to organize and retrieve learning content in geographically dispersed repositories. The main objective is to share and reuse data
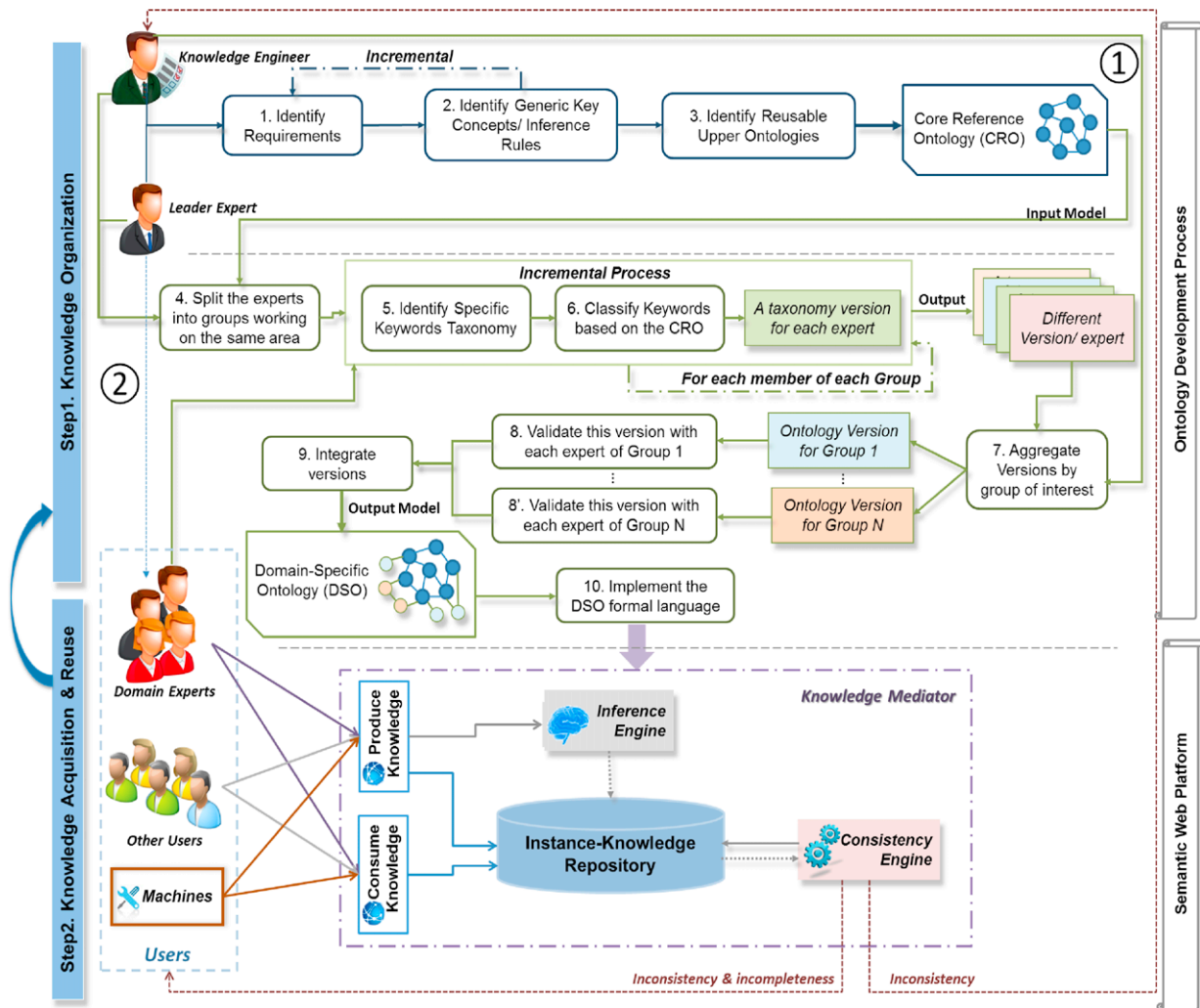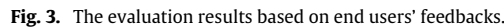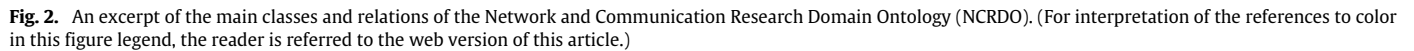

Fig. 1. The generic two-step methodology for formalizing experts' tacit knowledge.

grids to solve the learning content management problem and provide information retrieval technique that enhances the resource discovery mechanism of data grids based on ontology. Similar to the work of [21,22], we proposed a multi-level approach describing the knowledge to guarantee its modularity and reuse. However, none of [21,22] identified the methodology used to develop the ontology. Our work proposed a generic methodology dedicated to tacit knowledge conceptualization and formalization using semantic web. Another interesting work is the work of Li et al. [23] where the authors proposed an ontology-centric architecture, which focuses on project characterization, then, specialized it to the phenomic domain where new classes were added to generate specific domain ontology. However in [23], the authors did not formalize the experts' skills which are considered in our work.

None of the discussed works [21–23] clearly described how the knowledge is formalized and what methodology has been used to develop the ontology. In literature, many methodologies have been proposed for building ontology [24]. The most mature methodology is the METHONTOLOGY which was developed taking the IEEE 1074-1995 standard for software development process as a starting point [25]. It covers the ontology development process, the ontology life cycle and the techniques to achieve each activity [26] from the specification to the maintenance of the implemented ontology. Contrary to our proposed methodology, METHONTOLOGY does not provide information about how experts can be involved in this process to formalize their tacit knowledge. On-To-Knowledge [8] introduced the balance between human problem

solving and the automated IT solutions. It starts with a feasibility study where problems and people that can be involved are identified. Then, the ontology requirements are specified, a semi-formal ontology description is elaborated, and finally the applicative ontology is produced and evaluated. Similar to On-To-Knowledge, our methodology points out the importance of addressing human issues in the ontology development process. However, our work details the workflow to manage these issues and interactions with experts.

Both METHONTOLOGY and On-To-Knowledge do not consider collaborative construction of ontologies [27] which is mandatory to formalize the experts' tacit knowledge. In this context, Kotis and Vouros [28] proposed a Human-Centered Ontology Engineering Methodology (HCOME). Compared to existing methodologies [24], HCOME contributes to involving the knowledge worker who focuses on developing and managing his/her ontology in his/her personal space. These ontologies are shared with other workers for review and evaluation. Thus, this methodology requires skills in ontology modeling. Similar to HCOME, our methodology is based on experts collaboration, but relies on face-to-face exchange, which remains more effective and exhibit higher performance results [29], when identifying the meta-knowledge reflecting experts' vision. NeOn [30] is a scenario-based methodology that adopts a different vision from existing methodologies. It does not define workflow for the ontology development, but identifies nine scenarios for collaboratively building ontologies and ontology networks, reusing and re-engineering knowledge resources

Fig. 2. An excerpt of the main classes and relations of the Network and Communication Research Domain Ontology (NCRDO). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 3. The evaluation results based on end users' feedbacks.

(ontological and non-ontological). Our methodology integrates some scenarios provided by NeOn such as reusing existing ontological knowledge sources; but it deepens the workflow for building ontologies and copes with social challenges that can loom up the collaborative ontology construction such as conflicts and psychological barriers.

6. Conclusion

In this paper, we proposed a generic two-step methodology dedicated to collaboratively conceptualizing and formalizing the experts' tacit knowledge within a team. This methodology takes into consideration the human barriers and adopts a multi-level approach to generate the core ontology describing generic concepts, and the domain specific ontology portraying the experts'

skills. We presented the result of applying this methodology for managing the scientific research activities and we developed a semantic web platform to manage our team experts' tacit knowledge. The proposed methodology has been evaluated based on users' feedbacks and their quality of experience.

Currently, we are mainly working on automating the knowledge acquisition process by conceiving an intelligent information retrieval service that will contribute on the ontology-instance evolution. Our future work will focus on the knowledge maintenance and reasoning.

Acknowledgments

and Communication Research Domain Ontology and identify their interests.

References

[1] J.S. Katz, B.R. Martin, What is research collaboration? Res. Policy 26 (1997) 1–18.

[2] L. Tsui, S.A. Chapman, L. Schnirer, S. Stewart, A Handbook on Knowledge Sharing: Strategies and Recommendations for Researchers, Policy Makers and Service Providers, Community-University Partnership for the Study of Children, Youth, and Families Edmonton, 2006.

[3] C. Hussin, A. Razak, M. Hussain Alsayed, H. Mohamed Dahlan, Knowledge sharing in collaborative research activities, J. Inf. Syst. Res. Innov. 2 (2012) 1–8.

[4] R. Studer, V.R. Benjamins, D. Fensel, Knowledge engineering: principles and methods, Data Knowl. Eng. 25 (1998) 161–197.

[5] M. Earl, Knowledge management strategies: Toward a taxonomy, J. Manage. Inf. Syst. 18 (2001) 215–233.

[6] S.R. Ghani, Knowledge management: tools and techniques, DESIDOC J. Libr. Inf. Technol. 29 (2009) 33–38.

[7] R. Dieng, O. Corby, A. Giboin, M. Ribiere, Methods and tools for corporate knowledge management, Int. J. Hum.-Comput. Stud. 51 (1999) 567–598.

[8] Y. Sure, S. Staab, R. Studer, On-to-knowledge methodology (OTKM), in: Handbook on Ontologies, Springer, 2004, pp. 117–132.

[9] M. Alavi, D.E. Leidner, Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues, MIS Q. (2001) 107–136.

[10] S. Wang, R.A. Noe, Knowledge sharing: A review and directions for future research, Hum. Resour. Manag. Rev. 20 (2010) 115–131.

[11] C. Roussey, F. Pinet, M. Kang, O. Corcho, An introduction to ontologies and ontology engineering, in: Ontologies in Urban Development Projects, Springer, London, 2011, pp. 9–38.

[12] K. Yaakub, R. Shaari, S.A. Panatik, A. Rahman, Towards an understanding of the effect of core self-evaluations and knowledge sharing behaviour, Int. J. Appl. Psychol. 3 (2013) 13–18.

[13] L.A. Pervin, Personality: Theory, Assessment, and Research, 1970.

[14] A.T. Amayah, Knowledge sharing, personality traits and diversity: a literature review, in: Proceedings from The Midwest Research-to Practice Conference in Adult, Continuing, and Community Education. St. Louis, MO, USA, 2011.

[15] SRACO Ontology, Available at: http://homepages.laas.fr/emezghan/SRACO.owl.

[16] NCRDO Ontology, Available at: http://homepages.laas.fr/emezghan/NCRDO. owl.

[17] E. Mezghani, M. Da Silveira, C. Pruski, E. Exposito, K. Drira, A perspective of adaptation in healthcare, Stud. Health Technol. Inform. 205 (2014) 206–210.

[18] S. Schaffert, F. Bry, J. Baumeister, M. Kiesel, Semantic wikis, software, IEEE 25 (2008) 8–11.

[19] F. Bry, S. Schaffert, D. Vrandečić, K. Weiand, Semantic Wikis: Approaches, Applications, and Perspectives, Springer, 2012.

[20] M. Krötzsch, D. Vrandečić, M. Völkel, Semantic mediawiki, in: The Semantic Web-ISWC 2006, Springer, 2006, pp. 935–942.

[21] G. Pirró, C. Mastroianni, D. Talia, A framework for distributed knowledge management: Design and implementation, Future Gener. Comput. Syst. 26 (2010) 38–49.

[22] W.-C. Shih, C.-T. Yang, S.-S. Tseng, Ontology-based content organization and retrieval for SCORM-compliant teaching materials in data grids, Future Gener. Comput. Syst. 25 (2009) 687–694.

[23] Y.-F. Li, G. Kennedy, F. Ngoran, P. Wu, J. Hunter, An ontology-centric architecture for extensible scientific data management systems, Future Gener. Comput. Syst. 29 (2013) 641–653.

[24] Y. Sure, S. Staab, R. Studer, Ontology engineering methodology, in: Handbook on Ontologies, Springer, 2009, pp. 135–152.

[25] M. Fernández-López, Overview of methodologies for building ontologies, in: IJCAI99 Workshop on Ontologies and Problem-Solving Methods: Lessons Learned and Future Trends, Stockholm, 1999.

[26] M. Fernández-López, A. Gómez-Pérez, N. Juristo, METHONTOLOGY: From ontological art towards ontological engineering, in: AAAI Symposium on Ontological Engineering, Stanford, 1997.

[27] O. Corcho, M. Fernández-López, A. Gómez-Pérez, Methodologies, tools and languages for building ontologies. Where is their meeting point? Data Knowl. Eng. 46 (2003) 41–64.

[28] K. Kotis, G.A. Vouros, Human-centered ontology engineering: The HCOME methodology, Knowl. Inf. Syst. 10 (2006) 109–131.

[29] M.E. Warkentin, L. Sayeed, R. Hightower, Virtual teams versus face-to-face teams: An exploratory study of a web-based conference system*, Decis. Sci. 28 (1997) 975–996.

[30] M.C. Suarez-Figueroa, A. Gomez-Perez, M. Fernandez-Lopez, The NeOn methodology for ontology engineering, in: Ontology Engineering in a Networked World, Springer, 2012, pp. 9–34.

Emna Mezghani received the Engineering and M.S. degrees in Computer Science from the National School of Engineering of Sfax, Tunisia in September 2011 and September 2012 respectively. Currently, she is a Ph.D. student in Software Engineering at the University of Toulouse. Her research interests include collaborative systems, semantic web, knowledge management and big data.

Ernesto Exposito is an Associate Professor at INSA Toulouse and a researcher at the LAAS/CNRS laboratory in France. In 2004, he worked as Researcher in the National ICT Australia Limited (NICTA) Research Center in Sydney, Australia. His research interests include autonomic communication services aimed at satisfying the requirements of new generation distributed applications in heterogeneous networked environments. He has served as chairman and member for many Program Committees. He has participated in several European and French research projects. He is the author of more than 80 publications including international journals, regular and invited international conference papers, books and book chapters.

Khalil Drira, Ph.D. 1992 is Research Director, a full-time research position at the French National Center for Scientific Research (CNRS). His research interests include formal design, implementation, and provisioning of distributed communicating systems and cooperative networked services. His research activity addresses different topics in this field focusing on model-based analysis and design of correctness properties including robustness, adaptability and reconfiguration. He is or has been a member of the program committees of international and national conferences. He chaired different international conferences including IEEE International workshops on Collaboration Technologies: Infrastructure for Collaborative Enterprises (IEEE- WETICE) and the European Conference on Software Architecture (ECSA).