



HAL
open science

Recommending competent person in a digital ecosystem

Ning Wang, Marie-Hélène Abel, Jean-Paul Barthès, Elsa Negre

► **To cite this version:**

Ning Wang, Marie-Hélène Abel, Jean-Paul Barthès, Elsa Negre. Recommending competent person in a digital ecosystem. IEEE International Conference on Industrial Informatics and Computer Systems (CIICS 2016), Mar 2016, Sharjah, United Arab Emirates. pp.37-43, 10.1109/ICCSII.2016.7462436 . hal-01386798

HAL Id: hal-01386798

<https://hal.science/hal-01386798>

Submitted on 5 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Recommending Competent Person in a Digital Ecosystem

Ning Wang, Marie-Hélène Abel, Jean-Paul Barthès
Sorbonne Universités, Université de
Technologie de Compiègne, CNRS,
UMR 7253 Heudiasyc
Compiègne, France
Email: {ning.wang, marie-helene.abel, barthes}@utc.fr

Elsa Negre
Université Paris-Dauphine, CNRS,
UMR 7243 Lamsade
Paris, France
Email: elsa.negre@dauphine.fr

Abstract—When participating in activities in a collaborative digital environment, people leave traces. Such traces in return can be used for recommending somebody regarding his/her competency on a given concept. Doing so improves further collaboration because knowing users’ specialties ameliorates distributing tasks reasonably. In this article we propose a semantic model taking into account the concepts of activity, trace of interaction and competency. Applying Logistic Regression, we exploit this model to offer recommendation of a competent person in a digital ecosystem.

I. INTRODUCTION

A digital ecosystem is a self-organizing digital infrastructure aimed at creating a digital environment for networked organizations supporting cooperation, knowledge sharing, development of open and adaptive technologies and evolutionary business models, according to Uden et al. [1]. Wegner mentions in [2] that collaboration requires an interactive memory having two parts: (i) a combination of individual knowledge; and (ii) interpersonal awareness of others’ knowledge. People collaborate inside a digital ecosystem by participating in various activities. We are interested in recording and exploiting these activities and store them as traces. Analyzing the traces using a model can help measuring the competency of a user as shown by Rodrigues et al. [3]. Based on the information exploited from the traces, we can improve collaboration focusing on the reuse of traces for different purposes like decision aid and recommendation, as shown by Chang et al. [4].

Once people have useful knowledge, they are capable of transforming this knowledge for solving a problem or facing different situations as mentioned by Leblanc [5]. According to Beyou [6], competency is defined as the capacity to use knowledge efficiently. But measuring competency has always been difficult. We propose to measure it by using a model of competency applied to traces of interaction in a digital ecosystem. We consider actions like: *Who responds to which questions on which concept? Who consults resources of the concept? Who adds resources to the concept?* After answering the above questions, we conclude and recommend the most competent person.

The remainder of this paper is organized as follows. In Section II we identify various merits and limitations of some current approaches. Section III introduces the problem statement, previous work on semantic traces and our approach.

In Section IV we propose a prototype of collaborative platform “MEMORAE.” Section V discusses the advantages of our approach. Section VI concludes and mentions directions for future work.

II. RELATED WORK

Various articles have been published about exploiting the traces with the help of semantics. Sahay et al. [7] propose a novel conversational search and recommendation system that involves finding relevant information based on social interactions and feedback. Chen et al. [8] present a mechanism for personalized knowledge search and recommendation, adapting a suitable domain ontology according to the previous browsing and reading behavior of users. Breese et al. [9], Condliff et al. [10] and Pennock et al. [11] tried to provide recommendation using probabilistic models. According to the Rummel Model mentioned by Rothwell et al. [12], competency is defined as human performance. Different definitions of competency share three mutual fundamental items: resource, context and objective as remarked by Harzallah et al. [13]:

- A competency is constructed from resource categories, e.g. knowledge, know-how and capabilities or personal qualities.
- The competency context is related to the environment in which competency is situated.
- A competency is motivated by an objective. It is characterized by the acquisition of a goal or the accomplishment of one or several tasks.

Distinct from the mentioned approaches, we try to give recommendations using on semantic model of competency and applying Logistic Regression (see Cox [14]).

III. OUR APPROACH

In this section we present our previous work on modelling competency and the mathematical method we apply to measure it.

1) Previous Work on Model of Competency: In our previous work, the model of traces proposed by Li [15] allowed an elaborate analysis of interactions among users. We also proposed a model of competency in a collaborative context (Wang et al. [16]). Competency of a user is described by

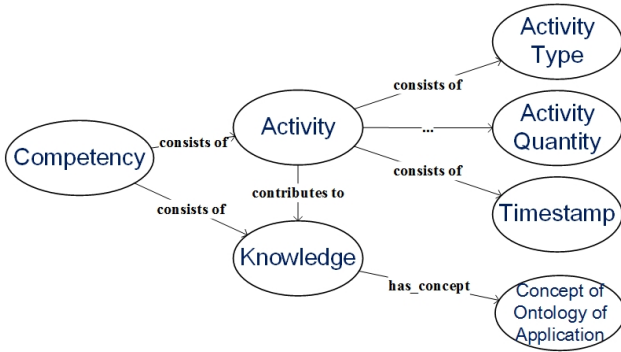


Fig. 1. Generic model of competency (Wang et al. [16]).

several components: an activity, an activity type, an activity quantity, a timestamp and a concept of ontology of application as shown in Figure 1.

The detailed definitions of the components are as follows:

- **Activity** denotes how a user applies knowledge. Activity also helps accumulating a user’s knowledge. For example, if a user consults many files about “Java,” it’s reasonable to assume that his/her knowledge grows on “Java.”
- **Activity Type** describes different types of activities. Some types of activities directly contribute to competency, for example answering questions from other users or creating a Wiki about a concept. Such activities indicate that the user tends to be more competent about what s/he applies. Meanwhile other activities only contribute to the knowledge about the concept like reading a paper about it.
- **Activity Quantity** records both a user and his collaborators’ intensity of efforts on this activity.
- **Timestamp** records the time when an activity takes place and terminates.
- **Knowledge** is what a user applies during an activity.
- **Concept of Ontology of Application** describes the nature of a user’s activity. It is the semantic description of knowledge.

2) *Logistic Regression*: Our approach consists of training a regressor based on various characters and traces user leaves in a digital ecosystem. Comparing with other regression algorithms that could be applied for our purpose, Logistic Regression as described by Cox [14] is adaptable to our case due to its high variability and non-linear distribution of input features. Whatever the input t is, output $H(t)$ is always restricted to the set $(0, 1)$. It uses the logistic function to model an output variable, namely:

$$H(t) = \exp(t)/(1 + \exp(t)) \quad (1)$$

A set of examples is represented by $(\mathbf{x}_i, v_i) \in R^n \times [0, 1]$ for the training set, i denoting a user, n being the number of features. In the training set, for each vector \mathbf{x}_i the corresponding value v_i is equal to 1 or 0 indicating whether \mathbf{x}_i belongs to a certain class or not. A set of weights $\mathbf{w} \in R^n$ is added,

TABLE I. AN EXAMPLE OF THREE DIFFERENT ACTIVITIES THAT USERS REALIZE

Act1: about_class: Cpt1; act_type: create; user: U1; indexKey: Ik1; date: 09-10-2015	Act2: about_class: Cpt1; act_type: access; user: U2; indexKey: Ik1; date: 10-10-2015	Act3: about_class: Cpt2; act_type: add; user: U1; indexKey: Ik2; date: 10-10-2015
-----------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------

TABLE II. COMPETENCY DEDUCTED FROM OBSERVATION OF ACTIVITIES IN TABLE I.

Competency1: owned_by: U1; about_class: Cpt1; evidence: createActivity; date: 09-10-2015	Competency2: owned_by: U2; about_class: Cpt1; evidence: accessActivity; date: 10-10-2015	Competency3: owned_by: U1; about_class: Cpt2; evidence: addActivity; date: 10-10-2015
------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------

determining the weight of each dimension of vector \mathbf{x}_i . We define the vector of features of user i on concept j as \mathbf{x}_i^j , a vector of dimension n . Each of the n dimensions corresponds to features concerning user’s performance in the platform. We will explain the detail of features in Section IV. Therefore (1) becomes:

$$H_w(\mathbf{x}_i^j) = \exp(\mathbf{w}^T \mathbf{x}_i^j) / (1 + \exp(\mathbf{w}^T \mathbf{x}_i^j)) \quad (2)$$

where $H_w(\mathbf{x}_i^j)$ equals to the probability that $v_i^j = 1$ given \mathbf{x}_i^j and \mathbf{w} .

$$H_{w,b}(\mathbf{x}_i^j) = P(v_i^j = 1 | \mathbf{x}_i^j, \mathbf{w}) \quad (3)$$

IV. FIRST PROTOTYPE

In this section, we present a prototype of MEMORAE developed by Abel et al. [17], a web based collaboration platform. First, we introduce the ontology of this platform. Then, using a toy example, we demonstrate how the recommender system works.

A. MEMORAE Collaborative Platform

MEMORAE is a web-based platform developed using web 2.0 technologies (Figure 2). It aims at facilitating knowledge sharing within organizations as claimed by Atrash et al. [18]. In this digital ecosystem, users organize themselves in networked organizations that support cooperation, knowledge sharing, the development of open and adaptive technologies and evolutionary business models. Different types of knowledge resources are supported: social (e.g. chat, event, wiki) or documentary. All types of knowledge resources (e.g. notes, documents, forums, etc.) are indexed by at least one concept of a semantic map. We added a Community-based Question Answering (CQA) service to this platform. When a question is proposed by a user, it is indexed by a concept in the map.

B. Extension of MEMORAE-core 2 model

Figure 3 displays the semantic model of MEMORAE. Competency-driven applications expect to have a clear statement that a user has a competency, but one of the key problems of competency-oriented approaches is how to diagnose competencies reliably. Usually, one can only observe the performance of a user and try to deduce from it the presence of a competence as mentioned by Kumuudha et al. [19]. We define the concept *Activity* in which a user participates as *CompetencyEvidence* as done by Schmidt and Kunzmann

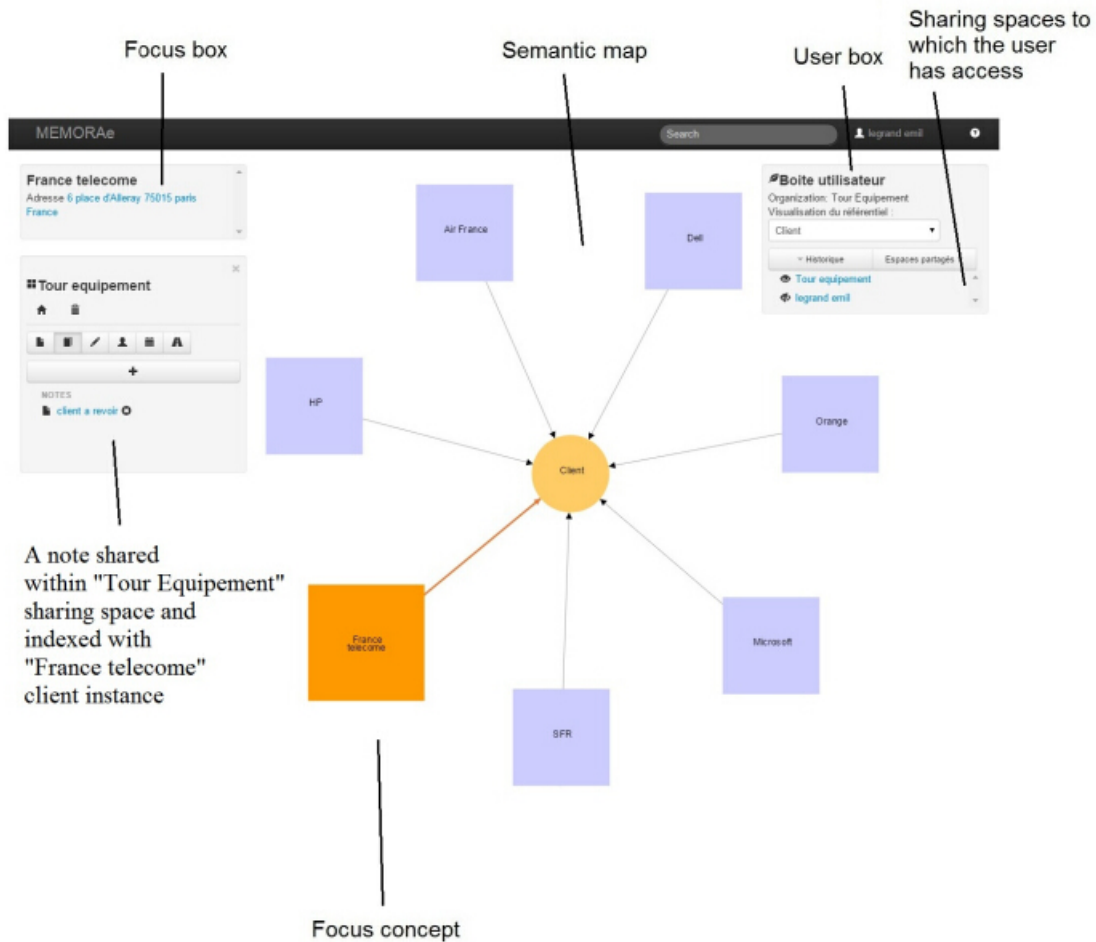


Fig. 2. Main interface of MEMORAe web application.

[20] to support competency. Combining with Figure 1, we have:

- **Activity** serves as evidence to diagnose competencies.
- **Activity** comes in three types: *createActivity*, *addActivity* and *accessActivity*.
- **Activity Quantity** counts the number of activities supporting competency.
- **Timestamp** is represented by the datatype property *xsd:dateTime*.
- **Knowledge** corresponds to the ontology of application (OWL Classes) to which competency is related.

We show how these properties are defined by an example. Suppose user U_1 creates a document on the concept C_{p1} indexed by Ik_1 on October 9, 2015. The next day, user U_2 get access to the same document. On the same day, user U_1 adds a document on C_{p2} indexed by Ik_2 . We represent these three activities respectfully Act1, Act2 and Act3 in Table I. Observing these activities, we deduce the competencies shown in Table II.

In Figure 3, we distinguish competency by three types according to the taxonomy of primitive competencies as done by Vasconcelos et al. [21]:

- **Cognitive Competency** contains intangible abilities which are not restricted to an ontology of application, e.g., comprehension, reasoning and creativity.
- **Project Competency** includes management and communication which are also not limited to a certain ontology of application.
- **Technical Competency** relies closely to concepts of a certain ontology of application. For example, a high level technical competency on mathematics can be proved by good ability on statistics.

A person has a role when participating in an activity. A role requires certain competencies, e.g., a professor must have a high level competency in his/her domain. In MEMORAe, a person participates in interactive activities in the environment by an online account. The more activities a person participates in, the more competency we believe s/he possesses. Interactive activities are distinguished by creation activity, access activity and add activity. Normally creation activity is more important than the other two types of activities. We register all activities a person participates in as a trace, in order to analyze his/her competency with the help of Logistic Regression.

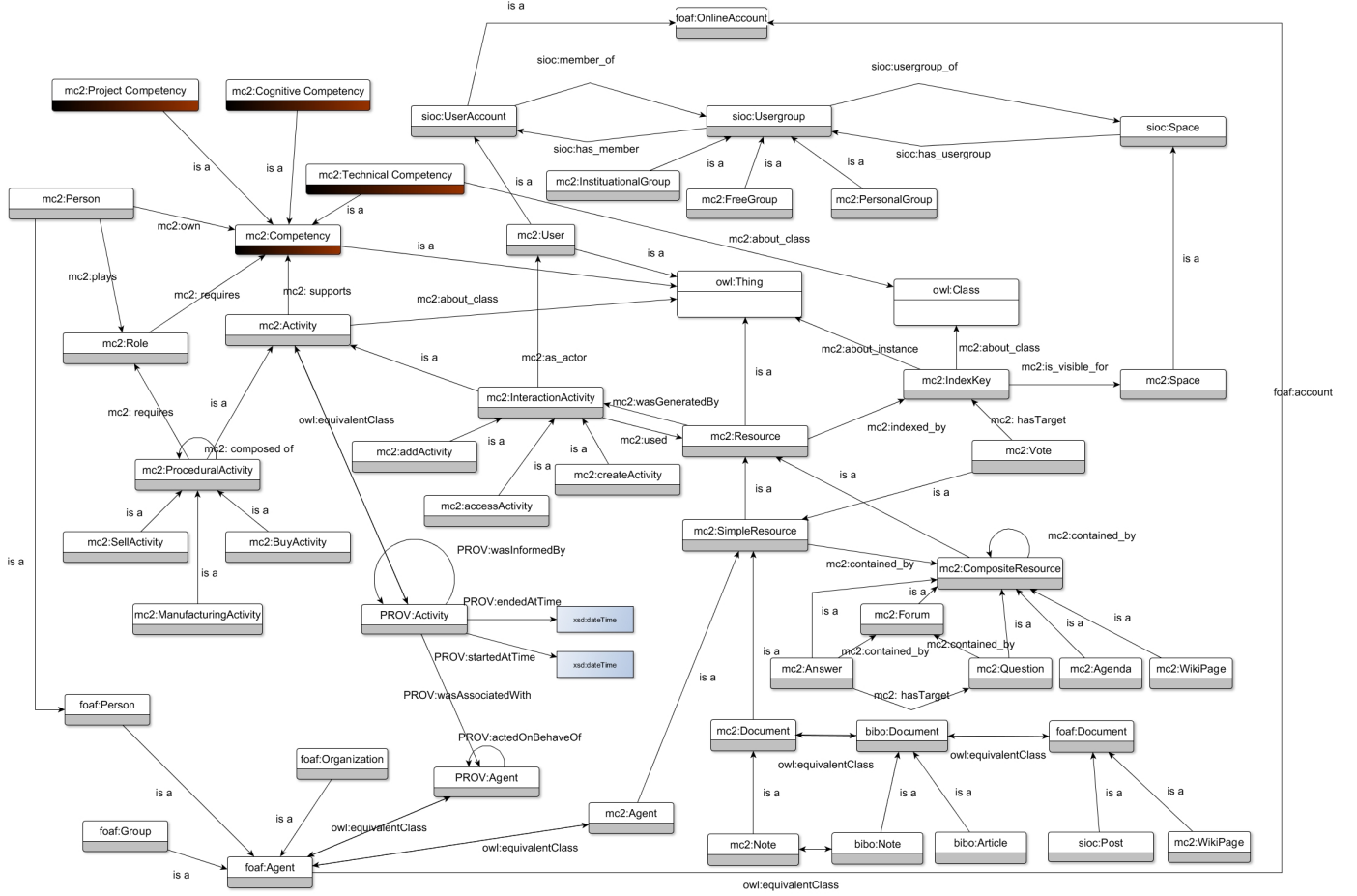


Fig. 3. Semantic model applied in MEMORAE (MEMORAE-core 2).

TABLE III. EXAMPLE OF TRAINING SET AND TESTING SETS

Training set						
User	Label	features				
		create	access	add	share	CLS2
A	0	4	55	6	2	0.30
B	1	5	70	4	8	0.70
C	0	2	28	10	3	0.20
D	1	5	56	8	12	0.78
E	0	4	5	6	15	0.68
F	0	2	46	17	3	0.21
G	0	2	56	27	6	0.22
Testing set						
1	v1	7	30	8	6	0.70
2	v2	6	75	5	1	0.82

C. Implementation of the Recommender System

For the vector of features \mathbf{x}_i^j competency of person i on concept j should be taken into account including all interactions in the platform, e.g., how many times i gets access to resources on j or how the resources s/he adds to the platform are rated (by votes) by members sharing the same space. We include $H(\mathbf{x}_i^k)$ the probability that i has the most competency on k to evaluate $H(\mathbf{x}_i^j)$, i.e., $H(\mathbf{x}_i^k) \in \mathbf{x}_i^j$. k and j are two concept that are semantically related.

As Figure 4, we introduce a Q&A Performance Set, User Profile and Traces in the platform as the training set $T(\mathbf{x}_i^j, v_{i,j})$ to train the Logistic Regression. To determine \mathbf{w} , we minimize

the squared loss as following:

$$L(\mathbf{w}) = \sum_{i=1}^m (H_w(\mathbf{x}_i^j) - v_i^j)^2 + \frac{\lambda}{2} \|\mathbf{w}\|^2 \quad (4)$$

where $\frac{\lambda}{2} \|\mathbf{w}\|^2$ is applied for regularization and λ controls the strength. Once the Logistic Regression model trained, we give recommendation on j by comparing $H(\mathbf{X}_i^j)$ with $x \in \{\text{members of the sharing space}\}$.

We illustrate the application of Logistic Regression by a fictitious example. Table III contains a training set including data of seven users and a testing set on a concept $C1$. In this data set, features *create*, *access*, *add* separately shows numbers of activities each user participated on the concept $C1$. The feature *votes* indicates the total votes received by the resources a user cites in the CQA service. concept $C2$ represents the probability of user being competent on concept $C2$ semantically close to $C1$. In the training set, users are distinguished by labels. A label of 1 means a competent user to recommend, while 0 is not. We substitute this training set into equation 4 and minimize the loss $L(\mathbf{w})$ to get a solution of parameter set \mathbf{w} . Back to equation 2, with the acquired \mathbf{w} and test set in Table III, we obtain $v_1 = 0.24$, $v_2 = 0.87$. It means user 2 has a probability of 0.87 to be a competent user. On the contrary, user 1 is not recommended with a probability of only 0.24.

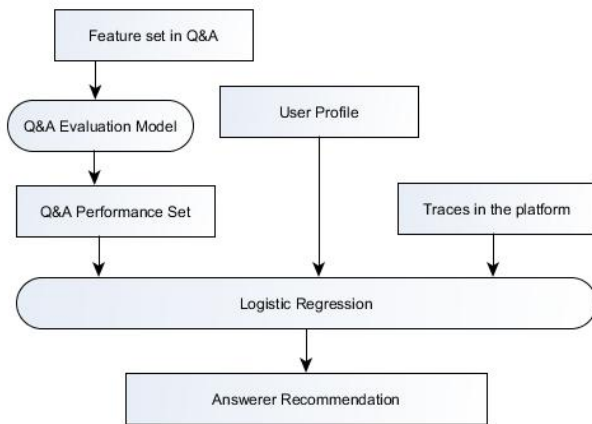


Fig. 4. The approach of recommender system based on features of CQA service and traces in the platform.

V. DISCUSSION

In previous section, we show how our approach works in evaluating user's competency. The general difficulty of evaluation is to decide the weight of each feature that affects the final result. For example, in our case we guess intuitively that creating a resource is more important than getting access to a resource on a same concept, but we have neither proof nor direct approach to quantify it. Thus we solve this challenge by three aspects. First, we distinguish activities into three different types as evidence to deduct competency so that different types have different weight. Second, how the resource users refer to is evaluated, is also considered as a feature in the evaluation. Third, users are not only evaluated by a single concept on which the newly arrived question is proposed, but also on the semantically related concepts.

VI. CONCLUSION AND FUTURE WORK

This paper presents an approach to valuate users competency given traces in the collaborative environment. To obtain this goal, we propose a model of competency. In a first prototype, we combine this model with the model of MEMORAE-core 2, a collaborative environment. We calculate users competency by applying Logistic Regression. Results show a good functionality to recommend competent person in this collaborative environment.

Further work includes collecting data from students attending a course on "methods and tools for knowledge and capitalization" at the University of Technology of Compiègne. During this course, students organize in MEMORAE to achieve their assigned watch and to produce at the end of the semester a report summarizing what they learnt from their investigations. At the end of this semester we will analyze each participant's competency in reference to their final score.

REFERENCES

[1] L. Uden, T. Wangsa, and E. Damiani, "The future of e-learning: E-learning ecosystem," in *Digital EcoSystems and Technologies Conference, 2007. DEST'07. Inaugural IEEE-IES*. IEEE, 2007, pp. 113–117.

[2] D. M. Wegner, "Transactive memory: A contemporary analysis of the group mind," in *Theories of group behavior*. Springer, 1987, pp. 185–208.

[3] J. A. R. Nt, J. M. Souza, G. B. Xexéo *et al.*, "Bringing knowledge into recommendation systems," in *Computer Supported Cooperative Work in Design (CSCWD), 2011 15th International Conference on*. IEEE, 2011, pp. 246–252.

[4] N. Chang, M. Irvan, and T. Terano, "A tv program recommender framework," *Procedia Computer Science*, vol. 22, pp. 561–570, 2013.

[5] A. Leblanc, "Environnement de collaboration et mémoire organisationnelle de formation dans un contexte d'apprentissage," Ph.D. dissertation, Compiègne, 2009.

[6] C. Beyou, *Manager les connaissances: [du knowledge management au développement des compétences dans l'organisation]*. Liaisons, 2003.

[7] S. Sahay, A. Venkatesh, and A. Ram, "Cobot: Real time multi user conversational search and recommendations," *Recommender Systems & the Social Web*, 2009.

[8] Y.-J. Chen, H.-C. Chu, Y.-M. Chen, and C.-Y. Chao, "Adapting domain ontology for personalized knowledge search and recommendation," *Information & Management*, vol. 50, no. 6, pp. 285–303, 2013.

[9] J. S. Breese, D. Heckerman, and C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering," in *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 1998, pp. 43–52.

[10] M. K. Condliff, D. D. Lewis, D. Madigan, and C. Posse, "Bayesian mixed-effects models for recommender systems," in *ACM SIGIR99 Workshop on Recommender Systems: Algorithms and Evaluation*, vol. 15, no. 5. Citeseer, 1999.

[11] D. M. Pennock, E. Horvitz, S. Lawrence, and C. L. Giles, "Collaborative filtering by personality diagnosis: A hybrid memory-and model-based approach," in *Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 2000, pp. 473–480.

[12] W. J. Rothwell and H. C. Kazanas, *Mastering the instructional design process: A systematic approach*. John Wiley & Sons, 2011.

[13] M. Harzallah and F. Vernadat, "It-based competency modeling and management: from theory to practice in enterprise engineering and operations," *Computers in industry*, vol. 48, no. 2, pp. 157–179, 2002.

[14] D. R. Cox, "The regression analysis of binary sequences," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 215–242, 1958.

[15] Q. Li, "Modeling and exploitation of the traces of interactions in the collaborative working environment," Ph.D. dissertation, Université de Technologie de Compiègne, 2013.

[16] N. Wang, M.-H. Abel, J.-P. Barthes, and E. Negre, "Mining user competency from semantic trace," in *Computer Supported Cooperative Work in Design (CSCWD), 2015 IEEE 19th International Conference on*. IEEE, 2015, pp. 48–53.

[17] M.-H. Abel, A. Benayache, D. Lenne, and C. Moulin, "E-memorae: a content-oriented environment for e-learning," in *E-Learning Networked Environments and Architectures*. Springer, 2007, pp. 186–205.

[18] A. Atrash, M.-H. Abel, and C. Moulin, "Notes and annotations as information resources in a social networking platform," *Computers in Human Behavior*, 2015.

[19] A. Kumuudha, K. Prabakar, and B. J. Godwin, "Organizational competency management: A competence performance approach," *International Journals of Marketing and Technology*, vol. 1, no. 7, pp. 116–121, 2011.

[20] A. Schmidt and C. Kunzmann, "Towards a human resource development ontology for combining competence management and technology-enhanced workplace learning," in *On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops*. Springer, 2006, pp. 1078–1087.

[21] J. Vasconcelos, C. Kimble, and A. Rocha, "Ontologies and the dynamics of organisational environments: an example of a group memory system for the management of group competencies," in *Proceedings of I-Know*, vol. 3, 2003, pp. 2–4.